

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:36:04 ; Search time 48 Seconds
(without alignments)
35.318 Million cell updates/sec

Title: US-09-912-414-2

Perfect score: 45

Sequence: 1 WVRWHF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

A_Geneseq_29Jan04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	6	3	AAB01492 Peptide w
2	41	91.1	6	3	AAB01505 Peptide w
3	41	91.1	256	3	AAG09463 Arabidops
4	41	91.1	256	3	AAG44075 Arabidops
5	41	91.1	256	3	AAG48293 Arabidops
6	41	91.1	325	3	AAG48292 Arabidops
7	41	91.1	325	3	AAG44074 Arabidops
8	41	91.1	325	3	AAG09462 Arabidops
9	41	91.1	325	7	ADB95074 A. thalia
10	40	88.9	306	3	AAG05810 Arabidops
11	40	88.9	306	3	AAG47194 Arabidops
12	40	88.9	338	3	AAG05809 Arabidops
13	40	88.9	338	3	AAG47193 Arabidops
14	40	88.9	418	2	AAG60501 Linoleic-
15	40	88.9	448	3	AAG47192 Arabidops
16	40	88.9	448	3	AAG05808 Arabidops
17	40	88.9	448	5	ABB93179 Herbicida
18	39	86.7	6	3	AAB01509 Peptide w
19	39	86.7	6	3	AAB01506 Peptide w
20	38	84.4	477	5	AAY51600 Human tru
21	38	84.4	477	5	AAB27932 Human tru
22	38	84.4	477	6	ABU56416 Lung canc
23	38	84.4	477	7	ABR82950 Human ort
24	38	84.4	537	5	AAE27933 Human Trk
25	38	84.4	537	7	ABR82951 Human Trk

26	38	84.4	553	5	AAE27934 Human Trk
27	38	84.4	553	7	ABR82952 Human Trk
28	38	84.4	822	2	AAE27931 Human Trk
29	38	84.4	822	3	AAE27931 Human Trk
30	38	84.4	822	5	AAE27931 Human Trk
31	38	84.4	822	5	AAE27931 Human Trk
32	38	84.4	822	5	AAE27931 Human Trk
33	38	84.4	822	7	ABR82949 Human Trk
34	38	84.4	822	7	ABR82949 Human Trk
35	38	84.4	822	7	ABR82949 Human Trk
36	36	80.0	79	3	ABP11225 Human ORF
37	36	80.0	79	5	ABP11225 Human ORF
38	36	80.0	83	2	AAE27931 Human sec
39	36	80.0	91	4	ABG07463 Novel hum
40	36	80.0	424	4	ABG13431 Soybean m
41	36	80.0	1169	4	ABG13431 Soybean m
42	35	77.3	6	3	AAE27931 Human sec
43	35	77.3	53	5	AAE27931 Human sec
44	35	77.3	53	5	AAE27931 Human sec
45	35	77.3	53	5	AAE27931 Human sec
46	35	77.3	53	5	AAE27931 Human sec
47	35	77.3	53	5	AAE27931 Human sec
48	35	77.3	53	5	AAE27931 Human sec
49	35	77.3	53	5	AAE27931 Human sec
50	35	77.3	53	5	AAE27931 Human sec
51	35	77.3	53	5	AAE27931 Human sec
52	34	75.6	6	3	AAE27931 Human sec
53	34	75.6	6	3	AAE27931 Human sec
54	34	75.6	6	3	AAE27931 Human sec
55	34	75.6	6	3	AAE27931 Human sec
56	34	75.6	6	3	AAE27931 Human sec
57	34	75.6	6	3	AAE27931 Human sec
58	34	75.6	6	3	AAE27931 Human sec
59	34	75.6	6	3	AAE27931 Human sec
60	34	75.6	6	3	AAE27931 Human sec
61	34	75.6	6	3	AAE27931 Human sec
62	34	75.6	6	3	AAE27931 Human sec
63	34	75.6	6	3	AAE27931 Human sec
64	34	75.6	6	3	AAE27931 Human sec
65	34	75.6	6	3	AAE27931 Human sec
66	34	75.6	6	3	AAE27931 Human sec
67	34	75.6	6	3	AAE27931 Human sec
68	34	75.6	6	3	AAE27931 Human sec
69	34	75.6	6	3	AAE27931 Human sec
70	34	75.6	6	3	AAE27931 Human sec
71	34	75.6	6	3	AAE27931 Human sec
72	34	75.6	6	3	AAE27931 Human sec
73	34	75.6	6	3	AAE27931 Human sec
74	34	75.6	6	3	AAE27931 Human sec
75	34	75.6	6	3	AAE27931 Human sec
76	34	75.6	6	3	AAE27931 Human sec
77	34	75.6	6	3	AAE27931 Human sec
78	34	75.6	6	3	AAE27931 Human sec
79	34	75.6	6	3	AAE27931 Human sec
80	34	75.6	6	3	AAE27931 Human sec
81	34	75.6	6	3	AAE27931 Human sec
82	34	75.6	6	3	AAE27931 Human sec
83	34	75.6	6	3	AAE27931 Human sec
84	34	75.6	6	3	AAE27931 Human sec
85	34	75.6	6	3	AAE27931 Human sec
86	34	75.6	6	3	AAE27931 Human sec
87	34	75.6	6	3	AAE27931 Human sec
88	34	75.6	6	3	AAE27931 Human sec
89	34	75.6	6	3	AAE27931 Human sec
90	34	75.6	6	3	AAE27931 Human sec
91	34	75.6	6	3	AAE27931 Human sec
92	34	75.6	6	3	AAE27931 Human sec
93	33	73.3	43	4	ABR16140 Human ner
94	33	73.3	43	4	ABR16140 Human ner
95	33	73.3	43	4	ABR16140 Human ner
96	33	73.3	43	4	ABR16140 Human ner
97	33	73.3	43	4	ABR16140 Human ner
98	33	73.3	43	4	ABR16140 Human ner

99	33	73.3	87	4	AAU56391	Aau56391 Propionib	172	33	73.3	1284	6	ABP57724	Abp57724 Protein #
100	33	73.3	87	6	ABM52910	Aam52910 Propionib	173	33	73.3	1603	4	AAU19518	Aau19518 Human dia
101	33	73.3	107	3	RAG03639	Agp03639 Human sec	174	33	73.3	1956	4	AAG65785	Agp65785 Human SNS
102	33	73.3	113	5	ABP64299	Abp64299 Human ORF	175	33	73.3	1956	4	AAG65785	Agp65785 Human SNS
103	33	73.3	119	5	ABP06139	Abp06139 Human ORF	176	33	73.3	1956	6	ABG76193	Abg76193 Human per
104	33	73.3	136	6	ABU01565	Abu01565 S. pneumo	177	33	73.3	1956	6	ABG75945	Abg75945 Human vol
105	33	73.3	137	3	AAU81510	Aay81510 Streptoco	178	33	73.3	1956	6	ABP72253	Abp72253 Human pn3
106	33	73.3	147	4	AAU90959	Aay90959 Human imm	179	33	73.3	1956	6	ADA50152	Ada50152 Human per
107	33	73.3	212	4	RAG91850	Agp91850 C glutami	180	33	73.3	1962	2	AAU17250	Aay17250 NaNG poly
108	33	73.3	214	6	ABU44730	Abu44730 Protein e	181	32	71.1	36	4	ABBA0893	Abb40893 Peptide #
109	33	73.3	227	3	AAG44583	Agp44583 Arabidops	182	32	71.1	36	4	AAU40893	Abb40893 Peptide #
110	33	73.3	227	3	RAG53102	Agp53102 Arabidops	183	32	71.1	36	4	AAU40893	Abb40893 Peptide #
111	33	73.3	232	5	ABP40375	Abp40375 Staphyloc	184	32	71.1	36	4	AAU40893	Abb40893 Peptide #
112	33	73.3	249	4	ABBS8731	Abbs8731 Drosophil	185	32	71.1	36	4	AAU40893	Abb40893 Peptide #
113	33	73.3	266	7	ADB74363	Adb74363 Mycobacte	186	32	71.1	46	2	AAU48521	Aay48521 Human bra
114	33	73.3	284	4	ABBS3008	Abbs3008 Escherich	187	32	71.1	51	6	ABP73090	Abp73090 Amino aci
115	33	73.3	286	3	RAG53101	Agp53101 Arabidops	188	32	71.1	52	6	ABP73090	Abp73090 Amino aci
116	33	73.3	286	3	AAG44582	Agp44582 Arabidops	189	32	71.1	52	6	ABP73090	Abp73090 Amino aci
117	33	73.3	289	7	ADC96004	Adc96004 E. faeciu	190	32	71.1	52	6	ABP73090	Abp73090 Amino aci
118	33	73.3	290	4	RAG81560	Agp81560 S. epider	191	32	71.1	59	5	ABP7602	Abp7602 Human ORF
119	33	73.3	296	4	ABG65116	Abg65116 Drosophil	192	32	71.1	60	3	AAU57596	Aau57596 Propionib
120	33	73.3	364	3	RAG44581	Agp44581 Arabidops	193	32	71.1	64	4	AAU57596	Aau57596 Propionib
121	33	73.3	365	3	RAG53100	Agp53100 Arabidops	194	32	71.1	64	4	AAU57596	Aau57596 Propionib
122	33	73.3	376	2	AAW46287	Aaw46287 Rhizobium	195	32	71.1	66	5	ABU67240	Abu67240 G-protein
123	33	73.3	377	2	AAW46288	Aaw46288 Brucella	196	32	71.1	66	5	ABU67240	Abu67240 G-protein
124	33	73.3	396	4	RAG98347	Agp98347 Escherich	197	32	71.1	68	5	ABP00898	Abp00898 Human ORF
125	33	73.3	396	6	ABU48344	Abu48344 Protein e	198	32	71.1	69	4	AAG75359	Agp75359 Human col
126	33	73.3	396	6	ABU14798	Abu14798 Protein e	199	32	71.1	78	4	ABP42209	Abp42209 Peptide #
127	33	73.3	402	6	ABJ25408	Abj25408 Aspergill	200	32	71.1	78	4	AAU36016	Aau36016 Peptide #
128	33	73.3	407	5	ABBS4596	Abbs4596 Lactococc	201	32	71.1	78	4	AAU36016	Aau36016 Peptide #
129	33	73.3	427	4	AAU20451	Aau20451 Human sec	202	32	71.1	78	4	AAU36016	Aau36016 Peptide #
130	33	73.3	427	4	AAU21839	Aau21839 Novel hum	203	32	71.1	78	4	AAU36016	Aau36016 Peptide #
131	33	73.3	441	6	ADC46480	Adc46480 Human neo	204	32	71.1	78	4	AAU36016	Aau36016 Peptide #
132	33	73.3	441	6	ABU38436	Abu38436 Protein e	205	32	71.1	80	5	ABP01776	Abp01776 Human ORF
133	33	73.3	454	5	ABG70036	Abg70036 Larval vi	206	32	71.1	80	5	ABP01776	Abp01776 Human ORF
134	33	73.3	456	4	AAU56511	Aau56511 Human pro	207	32	71.1	81	4	AAU15878	Aau15878 Peptide #
135	33	73.3	466	2	AAU56511	Aau56511 Human pro	208	32	71.1	81	4	AAU15878	Aau15878 Peptide #
136	33	73.3	476	5	AAE27936	Aae27936 TK negati	209	32	71.1	81	4	AAU15878	Aau15878 Peptide #
137	33	73.3	476	5	ABR82954	Abrr82954 Mouse tru	210	32	71.1	81	4	AAU15878	Aau15878 Peptide #
138	33	73.3	483	2	AAU68458	Aay68458 Mouse Trk	211	32	71.1	81	4	AAU15878	Aau15878 Peptide #
139	33	73.3	483	2	AAU44174	Aay44174 Bacillus	212	32	71.1	81	4	AAU15878	Aau15878 Peptide #
140	33	73.3	483	3	AAU78357	Aay78357 Bacillus	213	32	71.1	81	4	AAU15878	Aau15878 Peptide #
141	33	73.3	486	7	ABM74221	Abm74221 DNA clone	214	32	71.1	81	4	AAU15878	Aau15878 Peptide #
142	33	73.3	487	4	AAU42940	Aau42940 Propionib	215	32	71.1	81	4	AAU15878	Aau15878 Peptide #
143	33	73.3	487	6	ABM39459	Abm39459 Propionib	216	32	71.1	81	4	AAU15878	Aau15878 Peptide #
144	33	73.3	502	4	AAG90031	Agp90031 C. glutami	217	32	71.1	81	4	AAU15878	Aau15878 Peptide #
145	33	73.3	504	6	ABBS3206	Abbs3206 Protein s	218	32	71.1	84	3	AAU15878	Aau15878 Peptide #
146	33	73.3	504	6	ABU43208	Abu43208 Protein e	219	32	71.1	98	4	AAU15878	Aau15878 Peptide #
147	33	73.3	506	7	ADB70186	Adb70186 C. neofo	220	32	71.1	105	4	AAU15878	Aau15878 Peptide #
148	33	73.3	511	4	ABBS4917	Abbs4917 Drosophil	221	32	71.1	109	4	AAU15878	Aau15878 Peptide #
149	33	73.3	511	6	ABBS3204	Abbs3204 Protein s	222	32	71.1	109	4	AAU15878	Aau15878 Peptide #
150	33	73.3	516	4	AAU78876	Aay78876 C. glutam	223	32	71.1	109	4	AAU15878	Aau15878 Peptide #
151	33	73.3	516	5	ABP73570	Abp73570 Candida a	224	32	71.1	110	6	AAU15878	Aau15878 Peptide #
152	33	73.3	516	5	ABP73570	Abp73570 Candida a	225	32	71.1	113	6	AAU15878	Aau15878 Peptide #
153	33	73.3	529	3	AAU32200	Aay32200 Human rec	226	32	71.1	116	7	ABBS3886	Abbs3886 Human pro
154	33	73.3	553	4	ABG09100	Abg09100 Novel hum	227	32	71.1	123	5	ABBS3886	Abbs3886 Human pro
155	33	73.3	565	6	ABU49702	Abu49702 Protein e	228	32	71.1	123	5	ABBS3886	Abbs3886 Human pro
156	33	73.3	591	6	ABU26008	Abj26008 Aspergill	229	32	71.1	124	5	AAU78248	Abu78248 Polypepti
157	33	73.3	592	6	ABU03036	Abu03036 Moss lipi	230	32	71.1	126	2	AAU78248	Abu78248 Polypepti
158	33	73.3	708	4	ABG13429	Abg13429 Novel hum	231	32	71.1	126	2	AAU78248	Abu78248 Polypepti
159	33	73.3	751	4	ABG13429	Abg13429 Novel hum	232	32	71.1	129	5	ABG93128	Abg93128 S. cerevi
160	33	73.3	785	6	ABU39129	Abu39129 Protein e	233	32	71.1	132	6	ABP97359	Abp97359 Human ser
161	33	73.3	820	6	ABBS3886	Abbs3886 Phototrab	234	32	71.1	146	5	ABG93116	Abg93116 S. cerevi
162	33	73.3	821	5	AAE27935	Aae27935 Mouse ful	235	32	71.1	151	4	AAU78248	Abu78248 Polypepti
163	33	73.3	821	5	AAU50850	Aau50850 Rat recep	236	32	71.1	168	2	AAU78248	Abu78248 Polypepti
164	33	73.3	821	7	ADB79771	Adb79771 Rat neu	237	32	71.1	169	7	AAU78248	Abu78248 Polypepti
165	33	73.3	821	7	ABBS2953	Abbs2953 Mouse Trk	238	32	71.1	175	4	AAU78248	Abu78248 Polypepti
166	33	73.3	883	4	ABG29974	Abg29974 Novel hum	239	32	71.1	175	6	AAU78248	Abu78248 Polypepti
167	33	73.3	1014	3	ABBS15909	Abbs15909 E. coli p	240	32	71.1	177	7	AAU78248	Abu78248 Polypepti
168	33	73.3	1249	5	ABG30537	Abg30537 Alpha-iso	241	32	71.1	205	4	AAU78248	Abu78248 Polypepti
169	33	73.3	1284	5	ABG30563	Abg30563 Alpha-iso	242	32	71.1	217	4	AAU78248	Abu78248 Polypepti
170	33	73.3	1284	6	ABP70652	Abp70652 Bacillus	243	32	71.1	228	4	AAU78248	Abu78248 Polypepti
171	33	73.3	1284	6	ADA26479	Ada26479 Alpha-iso	244	32	71.1	244	4	AAU78248	Abu78248 Polypepti

245	32	71.1	32	318	71.1	484	6	ABU38941	Abu38941 Protein e
246	32	71.1	32	319	71.1	492	6	ABP78898	Abp78898 N. gonorr
247	32	71.1	32	320	71.1	492	6	ABP80291	Abp80291 N. gonorr
248	32	71.1	32	321	71.1	495	5	ABG61887	Abg61887 Prostate
249	32	71.1	32	322	71.1	495	6	ABU40750	Abu40750 Protein e
250	32	71.1	32	323	71.1	504	6	ABU37230	Abu37230 Protein e
251	32	71.1	32	324	71.1	504	6	ABU37658	Abu37658 Protein e
252	32	71.1	32	325	71.1	505	6	AAW06483	AAW06483 Helicobac
253	32	71.1	32	326	71.1	505	2	AAW06481	AAW06481 Helicobac
254	32	71.1	32	327	71.1	505	2	AAW06482	AAW06482 Helicobac
255	32	71.1	32	328	71.1	505	2	AAW54146	AAW54146 Helicobac
256	32	71.1	32	329	71.1	505	2	AAW52810	AAW52810 Helicobac
257	32	71.1	32	330	71.1	505	2	AAW98423	AAW98423 H. pylori
258	32	71.1	32	331	71.1	505	6	ABU30900	Abu30900 Protein e
259	32	71.1	32	332	71.1	506	6	ABU35323	Abu35323 Protein e
260	32	71.1	32	333	71.1	507	6	ABU42398	Abu42398 Protein e
261	32	71.1	32	334	71.1	508	2	AAU23340	AAU23340 A P450-1
262	32	71.1	32	335	71.1	508	3	ABU19693	Abu19693 Sweetgum
263	32	71.1	32	336	71.1	508	4	AAE05827	AAE05827 L. styrac
264	32	71.1	32	337	71.1	508	5	AAU77777	AAU77777 Arabidops
265	32	71.1	32	338	71.1	508	5	AAU77778	AAU77778 Modified
266	32	71.1	32	339	71.1	508	6	ABU30432	Abu30432 Protein e
267	32	71.1	32	340	71.1	509	2	AAU09191	AAU09191 Soybean c
268	32	71.1	32	341	71.1	517	3	AAW20650	AAW20650 Arabidops
269	32	71.1	32	342	71.1	517	3	AAW98268	AAW98268 H. pylori
270	32	71.1	32	343	71.1	527	3	AAW98268	AAW98268 H. pylori
271	32	71.1	32	344	71.1	567	3	AAW51001	AAW51001 Arabidops
272	32	71.1	32	345	71.1	570	5	ABR93717	ABR93717 Herbicida
273	32	71.1	32	346	71.1	599	6	ABP78908	Abp78908 N. gonorr
274	32	71.1	32	347	71.1	614	3	AAW51000	AAW51000 Arabidops
275	32	71.1	32	348	71.1	629	4	AAW70887	AAW70887 C albican
276	32	71.1	32	349	71.1	634	5	AAQ17252	AAQ17252 A thalian
277	32	71.1	32	350	71.1	697	3	AAW51020	AAW51020 Arabidops
278	32	71.1	32	351	71.1	709	3	AAW51019	AAW51019 Arabidops
279	32	71.1	32	352	71.1	721	5	ABR93758	ABR93758 Herbicida
280	32	71.1	32	353	71.1	728	3	AAW51018	AAW51018 Arabidops
281	32	71.1	32	354	71.1	739	5	ABR93759	ABR93759 Herbicida
282	32	71.1	32	355	71.1	871	5	ABR91828	ABR91828 Herbicida
283	32	71.1	32	356	71.1	878	5	ABR93298	ABR93298 Herbicida
284	32	71.1	32	357	71.1	1080	4	ABG25963	ABG25963 Novel hum
285	32	71.1	32	358	71.1	1285	1	AAW93089	AAW93089 Pasteurel
286	32	71.1	32	359	71.1	1508	4	AAW50676	AAW50676 C. elegan
287	32	71.1	32	360	71.1	1519	4	AAW50677	AAW50677 C. elegan
288	32	71.1	32	361	71.1	1527	4	ABR57771	ABR57771 Drosophil
289	32	71.1	32	362	71.1	4472	2	AAW97245	AAW97245 Virulence
290	32	71.1	32	363	71.1	956	6	ABU15390	ABU15390 Protein e
291	32	71.1	32	364	71.1	2858	4	ABR58064	ABR58064 Drosophil
292	32	71.1	32	365	71.1	3060	4	AAW58064	AAW58064 Drosophil
293	32	71.1	32	366	71.1	6	2	AAW37389	AAW37389 Peptide f
294	32	71.1	32	367	71.1	6	2	AAW28912	AAW28912 Oploid pe
295	32	71.1	32	368	71.1	6	2	AAW93770	AAW93770 New pepti
296	32	71.1	32	369	71.1	6	2	AAW23019	AAW23019 Oploid pe
297	32	71.1	32	370	71.1	6	3	AAW01507	AAW01507 Peptide w
298	32	71.1	32	371	71.1	7	4	AAW45777	AAW45777 H11 bindi
299	32	71.1	32	372	71.1	10	2	AAW86140	AAW86140 Anti-ELAM
300	32	71.1	32	373	71.1	10	2	AAW86145	AAW86145 Anti-ELAM
301	32	71.1	32	374	71.1	10	2	AAW86146	AAW86146 Anti-ELAM
302	32	71.1	32	375	71.1	10	2	AAW63963	AAW63963 ELAM-1 pe
303	32	71.1	32	376	71.1	10	2	AAW63964	AAW63964 ELAM-1 pe
304	32	71.1	32	377	71.1	10	2	AAW63958	AAW63958 ELAM-1 pe
305	32	71.1	32	378	71.1	19	4	AAW64620	AAW64620 Human sec
306	32	71.1	32	379	71.1	27	4	AAO12264	AAO12264 Human pol
307	32	71.1	32	380	71.1	30	4	AAO12524	AAO12524 Human pol
308	32	71.1	32	381	71.1	35	4	AAW67677	AAW67677 Human imm
309	32	71.1	32	382	71.1	36	6	ABU07779	ABU07779 Feline im
310	32	71.1	32	383	71.1	36	6	ABP98365	ABP98365 Amino aci
311	32	71.1	32	384	71.1	36	6	ABP98410	ABP98410 Amino aci
312	32	71.1	32	385	71.1	37	4	AAO13641	AAO13641 Human pol
313	32	71.1	32	386	71.1	41	4	AAO13323	AAO13323 Human pol
314	32	71.1	32	387	71.1	42	4	AAW33865	AAW33865 Peptide #
315	32	71.1	32	388	71.1	42	4	AAW73679	AAW73679 Human bon
316	32	71.1	32	389	71.1	42	4	ABG55415	ABG55415 Human liv
317	32	71.1	32	390	71.1	45	4	AAU17936	AAU17936 Novel hum
	32	71.1	32	391	71.1	49	4	AAW82540	AAW82540 Human imm

391	31	68.9	49	4	AAW74145	Human bon
392	31	68.9	49	4	AAO11866	Human pol
393	31	68.9	49	4	AAO13752	Human pol
394	31	68.9	49	4	ABG55923	Human liv
395	31	68.9	49	5	ABG44065	Human pep
396	31	68.9	50	4	ABBO3940	Human mus
397	31	68.9	50	6	ABU13234	Novel hum
398	31	68.9	51	4	AAW87511	Human inn
399	31	68.9	52	4	AAW90814	Human inn
400	31	68.9	52	4	AAU44331	Propionib
401	31	68.9	52	6	ABW40850	Propionib
402	31	68.9	54	6	ABP98351	Peptide d
403	31	68.9	55	4	AAW89555	Human inn
404	31	68.9	57	4	ABG15160	Novel hum
405	31	68.9	58	4	AAW86851	Human inn
406	31	68.9	59	5	ABP01137	Human ORF
407	31	68.9	60	4	AAO13834	Human pol
408	31	68.9	60	5	ABP09967	Human ORF
409	31	68.9	61	4	AAW90504	Human inn
410	31	68.9	64	4	AAW89458	Human inn
411	31	68.9	66	4	AAU20311	Human nov
412	31	68.9	68	4	AAW91757	Human inn
413	31	68.9	69	4	AAW88889	Human inn
414	31	68.9	69	4	AAO06382	Human pol
415	31	68.9	70	5	ABP06449	Human ORF
416	31	68.9	72	4	AAW80303	Human pro
417	31	68.9	72	4	AAO04191	Human pol
418	31	68.9	73	3	AAW51393	Human sec
419	31	68.9	73	6	ABU07758	Feline im
420	31	68.9	73	6	ABP98389	Peptide d
421	31	68.9	76	4	AAW94257	Human rep
422	31	68.9	77	3	AAW51465	Human sec
423	31	68.9	77	3	AAW51467	Human sec
424	31	68.9	77	4	AAO01224	Human gen
425	31	68.9	77	4	AAU86857	Novel hum
426	31	68.9	77	4	AAW92479	Human dig
427	31	68.9	77	4	AAU22524	Novel hum
428	31	68.9	77	7	ADB32364	Human nov
429	31	68.9	77	7	ADB60191	Connectiv
430	31	68.9	79	4	AAW90025	Human inn
431	31	68.9	80	4	AAW96687	Human rep
432	31	68.9	80	4	ABW96587	Human tes
433	31	68.9	81	4	ABG25557	Novel hum
434	31	68.9	85	3	AAW56462	Human pro
435	31	68.9	85	4	AAO08766	Human pol
436	31	68.9	86	7	ADB74493	Mycobacte
437	31	68.9	88	3	ABG60087	Arabidops
438	31	68.9	94	4	AAW82869	Human inn
439	31	68.9	99	4	AAW89959	Human inn
440	31	68.9	102	4	AAU42565	Propionib
441	31	68.9	102	6	ABW39084	Propionib
442	31	68.9	103	4	AAW23872	M pneumon
443	31	68.9	105	3	AAW42868	Human ORF
444	31	68.9	106	7	ADB64610	Human pro
445	31	68.9	109	4	ABG28332	Novel hum
446	31	68.9	111	5	AAO16123	Peroxidas
447	31	68.9	113	4	AAU49865	Propionib
448	31	68.9	113	6	ABW46384	Propionib
449	31	68.9	117	5	ABG71316	Human Sai
450	31	68.9	119	7	ADE72482	Human end
451	31	68.9	123	4	AAO04164	Human pol
452	31	68.9	127	4	AAU43535	Propionib
453	31	68.9	127	6	ABW40054	Propionib
454	31	68.9	128	4	AAW92661	Novel hum
455	31	68.9	128	4	AAU22655	Human nov
456	31	68.9	128	7	ADB32495	Human ORF
457	31	68.9	130	3	ABW40616	Human ORF
458	31	68.9	130	5</		

ALIGNMENTS

RESULT 1
AAB01492

AAB01492
 IR AAB01492 standard: peptide: 6 AA.

XX AAB01492;

XX DT 08-NOV-2000 (first entry)

DE peptide which binds to tra

XX
KW pNA binding: transcription factor: E2F; E2F-1; cell cycle; DP-1;

KW activation; transcription; proliferative disorder; psoriasis;
KW restenosis.

XX
CS
SyntheticXX
-302000675-

XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX XX

XX

XXXXXX

XX
Ft Mueller Ar, Kona Hawaii (K)

DR
WE1; 2000-332000/10;
XX

PI. REF. TOES PRINTING CO

PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.

PS Claim 6; Page 2; 42pp; English.

XX Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis

XX Sequence 6 AA;

Query Match 100.0%; Score 45; DB 3; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WVRWHF 6

Db 1 WVRWHF 6

RESULT 2

AAB01505

ID AAB01505 standard; peptide; 6 AA.

XX AAB01505;

DT 08-NOV-2000 (first entry)

DE Peptide which binds to transcription factor E2F-1 DNA binding domain.

XX DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;

KW activation; transcription; apoptosis; proliferative disorder; psoriasis;

KW restenosis.

OS Synthetic.

XX WO200044771-A1.

XX 03-AUG-2000.

XX 26-JAN-2000; 2000WO-GB000227.

XX 26-JAN-1999; 99GB-00001710.

XX (PROL-) PROLIFIX LTD.

XX Mueller R, Kontermann RE, Montigiani S;

XX WPI; 2000-532806/48.

XX Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.

FS Example; Page 26; 42pp; English.

XX Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis

XX Sequence 6 AA;

Query Match 91.1%; Score 41; DB 3; Length 6;

Best Local Similarity 83.3%; Pred. No. 1.4e+06;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WVRWHF 6

Db 1 WVRWHF 6

RESULT 3

AAG09463

ID AAG09463 standard; protein; 256 AA.

XX AAG09463;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 7409.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; Genetic mapping; Gene expression control; promoter;

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132563P.

XX 14-MAY-1999; 99US-0132566P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134219P.

XX 18-MAY-1999; 99US-0134370P.

XX 19-MAY-1999; 99US-0134768P.

XX 20-MAY-1999; 99US-0134941P.

XX 21-MAY-1999; 99US-0135124P.

XX 21-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

XX 25-MAY-1999; 99US-0136021P.

XX 27-MAY-1999; 99US-0136392P.

XX 28-MAY-1999; 99US-0136782P.

XX 01-JUN-1999; 99US-0137222P.

XX 03-JUN-1999; 99US-0137528P.

XX 04-JUN-1999; 99US-0137502P.

XX 07-JUN-1999; 99US-0137724P.

XX 08-JUN-1999; 99US-0138094P.

XX 10-JUN-1999; 99US-0138540P.

XX 10-JUN-1999; 99US-0138847P.

XX 14-JUN-1999; 99US-0139119P.

XX 18-JUN-1999; 99US-0139452P.

XX 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 13-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158029P.
PR 12-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.1%; Score 41; DB 3; Length 256;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
| | | | |
Db 72 WARWHF 77

RESULT 4

AAAG44075	18-JUN-1999;	99US-0139460P;
ID	AAAG44075 standard; protein; 256 AA.	99US-0139461P;
XX		99US-0139462P;
AC	AAAG44075;	99US-0139463P;
XX		99US-0139750P;
DT	18-OCT-2000 (first entry)	99US-0139763P;
XX		99US-0139817P;
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 55163.	99US-0139899P;
XX		99US-0140353P;
XX	Protein identification; signal transduction pathway; metabolic pathway;	99US-0140354P;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	99US-0140695P;
KW	termination sequence.	99US-0140823P;
XX		99US-0140991P;
OS	Arabidopsis thaliana.	99US-0141287P;
XX		99US-0141842P;
XX	EP1033405-A2.	99US-0142154P;
XX		99US-0142055P;
PD	06-SEP-2000.	99US-0142390P;
XX		99US-0142803P;
XX	25-FEB-2000; 2000EP-00301439.	99US-0142920P;
XX		99US-0142977P;
XX		99US-0143542P;
PR	25-FEB-1999; 99US-0121825P.	99US-0143542P;
PR	05-MAR-1999; 99US-0123180P.	99US-0143624P;
PR	09-MAR-1999; 99US-0123548P.	99US-0144005P;
PR	23-MAR-1999; 99US-0125788P.	99US-0144085P;
PR	25-MAR-1999; 99US-0126264P.	99US-0144086P;
PR	29-MAR-1999; 99US-0126785P.	99US-0144325P;
PR	01-APR-1999; 99US-0127462P.	99US-0144331P;
PR	06-APR-1999; 99US-0128234P.	99US-0144332P;
PR	08-APR-1999; 99US-0128714P.	99US-0144333P;
PR	16-APR-1999; 99US-0129845P.	99US-0144333P;
PR	19-APR-1999; 99US-0130077P.	99US-0144335P;
PR	21-APR-1999; 99US-0130449P.	99US-0144335P;
PR	23-APR-1999; 99US-0130510P.	99US-0144332P;
PR	28-APR-1999; 99US-0134449P.	99US-0144884P;
PR	30-APR-1999; 99US-0132048P.	99US-0145086P;
PR	04-MAY-1999; 99US-0132407P.	99US-0145088P;
PR	05-MAY-1999; 99US-0132484P.	99US-0145087P;
PR	06-MAY-1999; 99US-0132485P.	99US-0145089P;
PR	06-MAY-1999; 99US-0132487P.	99US-0145192P;
PR	07-MAY-1999; 99US-0132863P.	99US-0145145P;
PR	11-MAY-1999; 99US-0134256P.	99US-0145218P;
PR	14-MAY-1999; 99US-0134218P.	99US-0145224P;
PR	14-MAY-1999; 99US-0134219P.	99US-0145276P;
PR	14-MAY-1999; 99US-0134221P.	99US-0145913P;
PR	14-MAY-1999; 99US-0134370P.	99US-0145918P;
PR	18-MAY-1999; 99US-0134768P.	99US-0145919P;
PR	19-MAY-1999; 99US-0134941P.	99US-0145951P;
PR	20-MAY-1999; 99US-0135124P.	99US-0146386P;
PR	21-MAY-1999; 99US-0135353P.	99US-0146388P;
PR	24-MAY-1999; 99US-0135629P.	99US-0146389P;
PR	25-MAY-1999; 99US-0136021P.	99US-0147038P;
PR	27-MAY-1999; 99US-0136392P.	99US-0147204P;
PR	28-MAY-1999; 99US-0136782P.	99US-0147302P;
PR	01-JUN-1999; 99US-0137222P.	99US-0147192P;
PR	03-JUN-1999; 99US-0137528P.	99US-0147260P;
PR	04-JUN-1999; 99US-0137502P.	99US-0147303P;
PR	07-JUN-1999; 99US-0137724P.	99US-0147416P;
PR	08-JUN-1999; 99US-0138094P.	99US-0147493P;
PR	10-JUN-1999; 99US-0138540P.	99US-0147935P;
PR	10-JUN-1999; 99US-0138547P.	99US-0148171P;
PR	14-JUN-1999; 99US-0139119P.	99US-0148314P;
PR	16-JUN-1999; 99US-0139452P.	99US-0148341P;
PR	16-JUN-1999; 99US-0139453P.	99US-0148565P;
PR	17-JUN-1999; 99US-0139492P.	99US-0148684P;
PR	18-JUN-1999; 99US-0139454P.	99US-0149368P;
PR	18-JUN-1999; 99US-0139455P.	99US-0149175P;
PR	18-JUN-1999; 99US-0139456P.	99US-0149426P;
PR	18-JUN-1999; 99US-0139457P.	99US-0149722P;
PR	18-JUN-1999; 99US-0139458P.	99US-0149723P;
PR	18-JUN-1999; 99US-0139459P.	99US-0149929P;

```
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150894P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151086P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.1%; Score 41; DB 3; Length 256;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 72 WVRWHF 77

RESULT 5
AAG48293
ID AAG48293 standard; protein; 256 AA.
XX
AC AAG48293;
XX
XX 18-OCT-2000 (first entry)
XX
```

DE XX Arabidopsis thaliana protein fragment SEQ ID NO: 60969.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.

```
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.1%; Score 41; DB 3; Length 256;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWRF 6
Db 72 WVRWRF 77

RESULT 6
AAG48292
ID AAG48292 standard; protein; 325 AA.
XX
AC AAG48292;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60968.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
```

XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01231180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135623P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145224P.
PR 22-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.

```

PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157573P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.1%; Score 41; DB 3; Length 325;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
DB 141 WARWHF 146

RESULT 7
AAG4074
ID AAG44074 standard; protein; 325 AA.
XX
AC AAG44074;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55162.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EF1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX

```

```

PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131445P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.

```

```
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154033P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.

PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.1%; Score 41; DB 3; Length 325;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 141 WARWHF 146

RESULT 8
AAG09462
ID AAG09462 standard; protein; 325 AA.
XX AC AAG09462;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7408.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
```


PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144433P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158029P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.

```

PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159332P.
PR 14-OCT-1999; 99US-0159333P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.1%; Score 41; DB 3; Length 325;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WVRWHF 6
Db 141 WARWHF 146

RESULT 9
ADB95074
ID ADB95074 standard; protein; 325 AA.
XX
AC ADB95074;
XX
DT 04-DEC-2003 (first entry)
DE
DE A. thaliana protein 52949A #SEQ ID 72.
XX
KW Plant; herbicide; weed; crop field; growth; development.
XX
KW Arabidopsis thaliana.
XX
OS Arabidopsis thaliana.
XX
PN WO2003008440-A2.
XX
PD 30-JAN-2003.
XX
PF 16-JUL-2002; 2002WO-EP007929.
XX
PR 16-JUL-2001; 2001US-0305806P.
PR 20-FEB-2002; 2002US-0358416P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Levin JZ, Patton DA, Mcelver JA, Budziszewski GU, Zhou Q, Aux GW;
PI Tossberg J, Wegrich Glover L, Ashby CS, Thomas CR, Madhavan E;
PI Lewis S, Dunn J, Cates E, Law MD;
XX
XX WPI; 2003-229557/22.
DR N-PSDB; ADB95073.
XX
XX Identifying an herbicidal compound, useful for controlling undesirable
PT vegetation, comprises combining a polypeptide with a compound to be
PT tested for the ability to bind to the polypeptide or inhibit the activity
PT of the polypeptide.

```

```

XX Claim 4; SEQ ID NO 72; 273pp; English.
XX
CC The invention relates to a method for identifying a herbicidal compound.
CC The method of the invention comprises combining a polypeptide having at
CC least 90% identical to any one of 48 69-1008 residue amino acid sequences
CC (designated as P1-P48), given in the specification, with a compound to be
CC tested for the ability to bind to the polypeptide or inhibit the activity
CC of the polypeptide, under conditions conducive to binding or inhibiting,
CC respectively. Also disclosed is a method for killing or inhibiting the
CC growth or viability of a plant by applying to the plant the herbicidal
CC compound identified by the novel method, a chimeric construct comprising
CC a promoter operatively linked to the nucleic acid molecule, a recombinant
CC vector comprising the chimeric construct and a host cell comprising the
CC nucleic acid molecule. The method and polypeptides are useful in
CC screening assays to identify compounds that interact or inhibit the
CC polypeptides, thus as potential herbicides to control undesirable
CC vegetation such as weeds in crop fields. Nucleic acid molecules (odd
CC numbers between ADB95003 and ADB95097) isolated from Arabidopsis thaliana
CC comprising nucleotide sequences that encode proteins (even numbers
CC between ADB95004-ADB95098) are essential for plant growth and
CC development.
XX
XX Sequence 325 AA;
Qy Query Match 91.1%; Score 41; DB 7; Length 325;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 WVRWHF 6
Db 141 WARWHF 146

RESULT 10
AAG05810
ID AAG05810 standard; protein; 306 AA.
XX
AC AAG05810;
XX
DT 17-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2354.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125786P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.

```



```
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      88.9%; Score 40; DB 3; Length 306;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 WVRWHF 6
Db      115 WVNWHF 120

RESULT 11
ID AAG47194 standard; protein; 306 AA.
AC AAG47194;
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 59455.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0132864P.
XX 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
```

```
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149929P.
PR 25-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0158458P.
PR 29-SEP-1999; 99US-0158596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159295P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0160989P.
PR 26-OCT-1999; 99US-0161404P.
PR 26-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161539P.
PR 26-OCT-1999; 99US-0161539P.
PR 26-OCT-1999; 99US-0161616P.
PR 28-OCT-1999; 99US-0161616P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161993P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161539P.
PR 26-OCT-1999; 99US-0161539P.
PR 26-OCT-1999; 99US-0161616P.
PR 28-OCT-1999; 99US-0161616P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161993P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.
Query Match      88.9%; Score 40; DB 3; Length 306;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWEF 6
DB 115 WVRWEF 120
|||
|||
|||
RESULT 12
AAG05809
ID AAG05809 standard; protein; 338 AA.
AC AAG05809;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2353.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 200EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
```

PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139753P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140981P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144532P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 19-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.

```
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      88.9%; Score 40; DB 3; Length 338;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WVRWHF 6
Db      147 WVRWHF 152

RESULT 13
AAG47193
ID AAG47193 standard; protein; 338 AA.
AC AAG47193;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 59454.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135622P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140951P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
```

PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 28-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159594P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 88.9%; Score 40; DB 3; Length 338;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WYRWHF 6
DB 147 WYRWHF 152

RESULT 14

AAR60501
ID AAR60501 standard; protein; 418 AA.

XX AAR60501;
AC AAR60501;
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 29-MAR-1995 (first entry)
XX Linoleic-acid-desaturase.
XX Linoleic-acid-desaturase;
KW Linoleic-acid-desaturase; oleic-desaturase; transgenic plant;
XX crop improvement; linolenic acid.

XX Unidentified.

XX WO9418337-A1.

XX 18-AUG-1994.

XX 04-FEB-1994; 94WO-US001321.

XX 05-FEB-1993; 93US-00014431.

XX 22-NOV-1993; 93US-00156551.

XX (MONS) MONSANTO CO.

XX (UNMS) UNIV MICHIGAN STATE.

XX Gibson SI, Kishore GM, Ruff TG, Somerville CR, Aronel VJA;

XX WPI; 1994-279758/34.

XX N-PSDB; AAQ71242.

XX Genetically transformed plants with altered linolenic acid content -
PT contg recombinant, double-stranded DNA encoding linolenic acid
PT desaturase, or the antisense of the coding sequence.

XX Disclosure; Page 100-102; 144pp; English.

XX The isolation of cDNA encoding linoleic-acid-desaturase and oleic-

CC desaturase is described. A sequence of the invention is given in

CC AAQ71242, and its encoded protein sequence in AAR60501. (Updated on 25-

CC MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS

CC field.)

XX SQ Sequence 418 AA;

Query Match 88.9%; Score 40; DB 2; Length 418;

Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WYRWHF 6

DB 257 WYRWHF 262

RESULT 15

AAG47192
ID AAG47192 standard; protein; 448 AA.

XX AAG47192;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 59453.

[illegible]

```
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145227P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145911P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
```

```
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 88.9%; Score 40; DB 3; Length 448;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WVNWHF 6
Db 257 WVNWHF 262

RESULT 17
ABB93179
ID ABB93179 standard; protein; 448 AA.
XX AC
XX ABB93179;
XX 31-MAY-2002 (first entry)
DT
DE Herbicidally active polypeptide SEQ ID NO 2390.
XX
XX Herbicidal; plant; agriculture; herbicide.
XX OS Arabidopsis thaliana.
XX PN W0200210210-A2.
XX PD 07-FEB-2002.
XX
XX 28-AUG-2001; 2001WO-EP009892.
XX 28-AUG-2001; 2001WO-EP009892.
XX (FARB) BAYER AG.
XX

AC AAY51600;
 XX 30-MAY-2000 (first entry)
 DT Human truncated trkB receptor protein.
 DE
 XX trkB; human; receptor tyrosine kinase; trkB; diagnosis; neurotrophin;
 KW neurotrophic factor.
 XX Homo sapiens.
 OS
 XX US6027927-A.
 PN
 XX 22-FEB-2000.
 PD
 XX 01-OCT-1997; 97US-00942562.
 PF
 XX 18-MAR-1994; 94US-00215139.
 PR 05-AUG-1994; 94US-00286846.
 PR 19-MAY-1995; 95US-00444597.
 XX (GETH) GENENTECH INC.
 XX Urfer R, Shelton DL, Presta LG;
 PI WPI; 2000-194832/17.
 DR N-PSDB; AAZ88840.
 XX New human trk receptors useful in the diagnosis of various human
 PT pathological conditions associated with elevated or reduced levels of
 PT neurotrophins capable of binding trkB and/or trkC.
 XX Claim 5; Col 69-72; 78pp; English.
 PS This invention describes a novel isolated and purified polypeptide (I),
 CC belonging to the trk family of receptor tyrosine kinases, trkB and trkC.
 CC (I) are useful in the purification of human neurotrophic factors and in
 CC the diagnosis of various human pathological conditions associated with
 CC elevated or reduced levels of neurotrophins capable of binding trkB
 CC and/or trkC. This sequence represents a truncated form of the human trkB
 CC receptor described in the method of the invention
 XX
 SQ Sequence 477 AA;
 Query Match 84.4%; Score 38; DB 3; Length 477;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db |::|||
 4 WIRWH 8
 RESULT 21
 ID AAE27932 standard; protein; 477 AA.
 AC AAE27932;
 XX 27-DEC-2002 (first entry)
 DT Human truncated TrkB (TrkB.T1) protein.
 DE
 XX Human; neurodegenerative disorder; neurodevelopmental disorder; TrkB;
 KW trkB; Alzheimer's disease; AD; amyotrophic lateral sclerosis; ALS; PD;
 KW Parkinson's disease; Huntington's disease; HD; Lou Gehrig's disease;
 KW diabetic peripheral neuropathy; Down's syndrome; DS; neuroprotective;
 KW gene therapy; anticonvulsant; cerebroprotective; nootropic.
 XX Homo sapiens.
 OS
 XX WO200267858-A2.
 PN
 XX

PD 06-SEP-2002.
 XX 22-FEB-2002; 2002WO-US005151.
 PF 22-FEB-2001; 2001US-0270553P.
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA Krueger BK, Kingsbury TJ, Bambrick LL, Dorsey SG;
 XX WPI; 2002-698627/75.
 DR N-PSDB; AAD45787.
 XX Treating and/or preventing neurodegenerative and neurodevelopmental
 PT disorders such as Alzheimer's, Parkinson's and Huntington's diseases by
 PT altering the ratio of amount of full-length and truncated TrkB or TrkC
 PT polypeptides.
 XX Disclosure; Page 61-63; 96pp; English.
 PS The present invention relates to a method of treating neurodegenerative
 CC or neurodevelopmental disorders in a mammal which involves administering
 CC an isolated nucleic acid encoding a full-length TrkB or TrkC or their
 CC mutant, variant, homologue or fragment or an anti-sense RNA for truncated
 CC TrkB or TrkC isoforms, where they increase the amount of full-length TrkB
 CC or TrkC or decrease the amount of truncated TrkB or TrkC in treated
 CC neurons. The methods and compositions of the invention are useful for
 CC treating or preventing neurodegenerative or neurodevelopmental disorders
 CC such as Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's
 CC disease (HD), amyotrophic lateral sclerosis (ALS; Lou Gehrig's disease),
 CC diabetic peripheral neuropathy, the adverse complications of Down's
 CC syndrome (DS) and other types of peripheral neuropathy. Sequences of the
 CC invention are also used in gene therapy. The present sequence is human
 CC truncated TrkB (TrkB.T1) protein
 XX
 SQ Sequence 477 AA;
 Query Match 84.4%; Score 38; DB 5; Length 477;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db |::|||
 4 WIRWH 8
 RESULT 22
 ID ABUS6416 standard; protein; 477 AA.
 XX ABUS6416;
 AC ABUS6416;
 XX 02-APR-2003 (first entry)
 DT Lung cancer-associated polypeptide #9.
 DE
 XX Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 OS
 XX WO200286443-A2.
 PN 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US012476.
 PF 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR

PI Krueger BK, Kingsbury TJ, Bambrick LL, Dorsey SG;
 XX WPI; 2002-698627/75.
 DR N-PSDB; AAD45788.
 XX
 PT Treating and/or preventing neurodegenerative and neurodevelopmental
 PT disorders such as Alzheimer's, Parkinson's and Huntington's diseases by
 PT altering the ratio of amount of full-length and truncated TrkB or TrkC
 PT polypeptides.
 XX
 PS Disclosure; Page 67-70; 96pp; English.
 XX
 CC The present invention relates to a method of treating neurodegenerative
 CC or neurodevelopmental disorders in a mammal which involves administering
 CC an isolated nucleic acid encoding a full-length TrkB or TrkC or their
 CC mutant, variant, homologue or fragment or an anti-sense RNA for truncated
 CC TrkB or TrkC isoforms, where they increase the amount of full-length TrkB
 CC or TrkC or decrease the amount of truncated TrkB or TrkC in treated
 CC neurons. The methods and compositions of the invention are useful for
 CC treating or preventing neurodegenerative or neurodevelopmental disorders
 CC such as Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's
 CC disease (HD), amyotrophic lateral sclerosis (ALS), Lou Gehrig's disease,
 CC diabetic peripheral neuropathy, the adverse complications of Down's
 CC syndrome (DS) and other types of peripheral neuropathy. Sequences of the
 CC invention are also used in gene therapy. The present sequence is human
 CC TrkB.Shc isoform protein
 XX
 SQ Sequence 537 AA;
 Query Match 84.4%; Score 38; DB 5; Length 537;
 Best Local Similarity 80.0%; Pred. No. 4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db |::|
 4 WIRWH 8
 RESULT 25
 ABR82951
 ID ABR82951 standard; protein; 537 AA.
 XX
 AC ABR82951;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human TrkB isoform TrkB.Shc isoform 1 polypeptide.
 XX
 KW TrkB; TrkC; neurodegenerative; neuro-developmental; antiparkinsonian;
 KW neuroprotective; anticonvulsant; cerebroprotective; vasotropic; isoform;
 KW nootropic; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003071872-A1.
 XX
 PD 04-SEP-2003.
 XX
 PF 28-MAY-2002; 2002WO-US016807.
 XX
 PR 22-FEB-2002; 2002WO-US005151.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 PA (KRUEG/) KRUEGER B K.
 PA (KING/) KINGSBURY T J.
 PA (BAMB/) BAMBRICK L L.
 PA (DORS/) DORSEY S G.
 XX
 PI Krueger BK, Kingsbury TJ, Bambrick LL, Dorsey SG;
 XX WPI; 2003-731549/69.
 DR N-PSDB; ACF36567.
 XX

PT Treating and/or preventing neurodegenerative or neuro-developmental
 PT disorders, such as Alzheimer's disease, Parkinson's disease and
 PT amyotrophic lateral sclerosis, using nucleic acids encoding TrkB and/or
 PT TrkC polypeptides.
 XX
 PS Disclosure; Page 69-71; 99pp; English.
 XX
 CC The invention relates to treating a neurodegenerative or neuro-
 CC developmental disorder in a mammal. The method involves altering the
 CC ratio of the amount of full length TrkB polypeptide to the amount of
 CC truncated TrkB polypeptides in a neuron or by altering the ratio of the
 CC amount of full length TrkC polypeptide to the amount of truncated TrkC
 CC polypeptides in a neuron. The methods and compositions of the present
 CC invention are useful for treating and/or preventing a neurodegenerative
 CC or neuro-developmental disorder, such as Alzheimer's disease, Parkinson's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis (Lou
 CC Gehrig's disease), the adverse neurologic complications of Down syndrome,
 CC diabetic peripheral neuropathy and other types of peripheral neuropathy,
 CC and is associated with an injury to the central or peripheral nervous
 CC system resulting from stroke, cerebral ischaemia, or chemical and/or
 CC physical trauma. The present sequence represents a human TrkB isoform
 CC TrkB.Shc isoform 1 (GenBank Accession No. AF410900)
 XX
 SQ Sequence 537 AA;
 Query Match 84.4%; Score 38; DB 7; Length 537;
 Best Local Similarity 80.0%; Pred. No. 4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db |::|
 4 WIRWH 8
 RESULT 26
 AAE27934
 ID AAE27934 standard; protein; 553 AA.
 XX
 AC AAE27934;
 XX
 DT 27-DEC-2002 (first entry)
 XX
 DE Human TrkB.Shc isoform protein #2.
 XX
 KW Human; neurodegenerative disorder; neurodevelopmental disorder; TrkB;
 KW TrkC; Alzheimer's disease; AD; amyotrophic lateral sclerosis; ALS; PD;
 KW Parkinson's disease; Huntington's disease; HD; Lou Gehrig's disease;
 KW diabetic peripheral neuropathy; Down's syndrome; DS; neuroprotective;
 KW gene therapy; anticonvulsant; cerebroprotective; nootropic.
 XX
 OS Homo sapiens.
 XX
 PN WO200267858-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 22-FEB-2002; 2002WO-US005151.
 XX
 PR 22-FEB-2001; 2001US-0270553P.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Krueger BK, Kingsbury TJ, Bambrick LL, Dorsey SG;
 XX WPI; 2002-698627/75.
 DR N-PSDB; AAD45789.
 XX
 PT Treating and/or preventing neurodegenerative and neurodevelopmental
 PT disorders such as Alzheimer's, Parkinson's and Huntington's diseases by
 PT altering the ratio of amount of full-length and truncated TrkB or TrkC
 PT polypeptides.
 XX
 PS Disclosure; Page 74-77; 96pp; English.

XX The present invention relates to a method of treating neurodegenerative
 CC or neurodevelopmental disorders in a mammal which involves administering
 CC an isolated nucleic acid encoding a full-length TrkB or TrkC or their
 CC mutant, variant, homologue or fragment or an anti-sense RNA for truncated
 CC TrkB or TrkC isoforms, where they increase the amount of full-length TrkB
 CC or TrkC or decrease the amount of truncated TrkB or TrkC in treated
 CC neurons. The methods and compositions of the invention are useful for
 CC treating or preventing neurodegenerative or neurodevelopmental disorders
 CC such as Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's
 CC disease (HD), amyotrophic lateral sclerosis (ALS; Lou Gehrig's disease),
 CC diabetic peripheral neuropathy, the adverse complications of Down's
 CC syndrome (DS) and other types of peripheral neuropathy. Sequences of the
 CC invention are also used in gene therapy. The present sequence is human
 CC TrkB.shc isoform protein

XX Sequence 553 AA;
 SQ Query Match 84.4%; Score 38; DB 5; Length 553;
 Best Local Similarity 80.0%; Pred. No. 4.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 Db |:| | |
 4 WIRWH 8

RESULT 27
 ABR82952
 ID ABR82952 standard; protein; 553 AA.
 XX
 AC ABR82952;
 DT 18-DEC-2003 (first entry)
 XX
 DE Human TrkB isoform TrkB.Shc isoform 2 polypeptide.

XX TrkB; TrkC; neurodegenerative; neuro-developmental; antiparkinsonian;
 KW neuroprotective; anticonvulsant; cerebroprotective; vasotropic; isoform;
 KW nootropic; human.
 XX Homo sapiens.
 OS
 XX WO2003071872-A1.
 PN
 XX 04-SEP-2003.
 PD
 XX 28-MAY-2002; 2002WO-US016807.
 PF
 XX 22-FEB-2002; 2002WO-US005151.
 PR
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA (KRUE/) KRUEGER B K.
 PA (KING/) KINGSBURY T J.
 PA (BAMB/) BAMBRICK L L.
 PA (DORS/) DORSEY S G.
 XX
 XX Krueger BK, Kingsbury TJ, Bambrick LL, Dorsey SG;
 PI
 XX WPI; 2003-731549/69.
 DR N-PSDB; ACF36568.
 DR
 XX Treating and/or preventing neurodegenerative or neuro-developmental
 CC disorders, such as Alzheimer's disease, Parkinson's disease and
 CC amyotrophic lateral sclerosis, using nucleic acids encoding TrkB and/or
 CC TrkC polypeptides.
 PT
 XX Disclosure; Page 75-78; 99pp; English.

XX The invention relates to treating a neurodegenerative or neuro-
 CC developmental disorder in a mammal. The method involves altering the
 CC ratio of the amount of full length TrkB polypeptide to the amount of
 CC truncated TrkB polypeptides in a neuron or by altering the ratio of the

CC amount of full length TrkC polypeptide to the amount of truncated TrkC
 CC polypeptides in a neuron. The methods and compositions of the present
 CC invention are useful for treating and/or preventing a neurodegenerative
 CC or neuro-developmental disorder, such as Alzheimer's disease, Parkinson's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis (Lou
 CC Gehrig's disease), the adverse neurologic complications of Down syndrome,
 CC diabetic peripheral neuropathy and other types of peripheral neuropathy,
 CC and is associated with an injury to the central or peripheral nervous
 CC system resulting from stroke, cerebral ischaemia, or chemical and/or
 CC physical trauma. The present sequence represents a human TrkB isoform
 CC TrkB.Shc isoform 2 (GenBank Accession No. AF410901)

XX Sequence 553 AA;
 SQ Query Match 84.4%; Score 38; DB 7; Length 553;
 Best Local Similarity 80.0%; Pred. No. 4.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 Db |:| | |
 4 WIRWH 8

RESULT 28
 AAR81630
 ID AAR81630 standard; protein; 822 AA.
 XX
 AC AAR81630;
 DT 31-MAR-1996 (first entry)
 XX
 DE Human trkB receptor protein.

XX trkB receptor; tyrosine-kinase; enzyme; protease; inflammation; pain;
 KW diagnosis; neurotrophic factor; kidney; lung; psychiatric disorder.
 KW Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 67. .70
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 95. .98
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 121. .124
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 178. .181
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 205. .208
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 241. .244
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 254. .257
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 280. .283
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 325. .328
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 338. .341
 FT /note= "potential N-linked glycosylation site"
 FT Misc-difference 412. .415
 FT /note= "potential N-linked glycosylation site"
 FT Domain 431. .454
 FT /note= "transmembrane domain"
 FT Misc-difference 466
 FT /note= "splice site for truncated trkB"
 FT Domain 544. .807
 FT /note= "tyrosine-kinase domain"

XX WO9525795-A1.
 PN
 XX 28-SEP-1995.
 PD
 XX 17-MAR-1995; 95WO-US003426.
 PF

XX 18-MAR-1994; 94US-00215139.
 PR 05-AUG-1994; 94US-00286846.
 PR 20-DEC-1994; 94US-00359705.
 XX (GETH) GENENTECH INC.
 XX Presta LG, Shelton DL, Urfer R;
 XX WPI; 1995-344616/44.
 DR N-PSDB; AAT00689.
 XX New human trkB and trkC poly.peptide(s) and fusion proteins contg. them -
 PT also DNA, vectors and transformed cells useful in treatment and diagnosis
 PT of abnormal neurotrophic factor expression, e.g. inflammatory pain.
 XX Claim 8; Fig 1A-B; 117pp; English.
 XX This sequence may be expressed recombinantly for the production of human
 CC trkB receptor, and to detect or amplify trkB genes. The encoded protein
 CC may be used as a reagent in kinase receptor activation assays, and
 CC therapeutically in diseases associated with over or under expression of
 CC neurotrophic factor (e.g. pain of inflammation, kidney, lung,
 CC cardiovascular or psychiatric disorders and some sorts of tumours). At
 CC the indicated splice site, AA 436- 791 may be replaced by the sequence
 CC FVLPHKIPLDG (AAR81624)
 XX FVLPHKIPLDG (AAR81624)
 XX Sequence 822 AA;
 SQ
 Query Match 84.4%; Score 38; DB 2; Length 822;
 Best Local Similarity 80.0%; Pred. No. 6.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db |::|||
 4 WIRWH 8
 RESULT 29
 AAY51599
 ID AAY51599 standard; protein; 822 AA.
 AC AAY51599;
 XX 30-MAY-2000 (first entry)
 DT Human trkB receptor protein.
 DE trkB; human; receptor tyrosine kinase; trkC; diagnosis; neurotrophin;
 KW neurotrophic factor.
 KW Homo sapiens.
 XX US6027927-A.
 PN 22-FEB-2000.
 PD 01-OCT-1997; 97US-00942562.
 XX 18-MAR-1994; 94US-00215139.
 PR 05-AUG-1994; 94US-00286846.
 PR 19-MAY-1995; 95US-00444597.
 XX (GETH) GENENTECH INC.
 XX Urfer R, Shelton DL, Presta LG;
 XX WPI; 2000-194832/17.
 DR N-PSDB; AAZ88839.
 XX New human trk receptors useful in the diagnosis of various human
 PT pathological conditions associated with elevated or reduced levels of
 PT neurotrophins capable of binding trkB and/or trkC.

XX Claim 1; Col 63-68; 78pp; English.
 XX This invention describes a novel isolated and purified polypeptide (I),
 CC belonging to the trk family of receptor tyrosine kinases, trkB and trkC.
 CC (I) are useful in the purification of human neurotrophic factors and in
 CC the diagnosis of various human pathological conditions associated with
 CC elevated or reduced levels of neurotrophins capable of binding trkB
 CC and/or trkC. This sequence represents the human trkB receptor described
 CC in the method of the invention
 XX Sequence 822 AA;
 SQ
 Query Match 84.4%; Score 38; DB 3; Length 822;
 Best Local Similarity 80.0%; Pred. No. 6.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db |::|||
 4 WIRWH 8
 RESULT 30
 AAE27931
 ID AAE27931 standard; protein; 822 AA.
 XX AAE27931;
 AC AAE27931;
 XX 27-DEC-2002 (first entry)
 DT Human trkB protein.
 DE Human; neurodegenerative disorder; neurodevelopmental disorder; TrkB;
 KW TrkC; Alzheimer's disease; AD; amyotrophic lateral sclerosis; ALS; PD;
 KW Parkinson's disease; Huntington's disease; HD; Lou Gehrig's disease;
 KW diabetic peripheral neuropathy; Down's syndrome; DS; neuroprotective;
 KW gene therapy; anticonvulsant; cerebroprotective; nootropic.
 XX Homo sapiens.
 OS WO200267858-A2.
 XX 06-SEP-2002.
 PD 22-FEB-2002; 2002WO-US005151.
 XX 22-FEB-2001; 2001US-0270553P.
 PR (UYMA-) UNIV MARYLAND BALTIMORE.
 XX Krueger BK, Kingsbury TJ, Bambrick LL, Dorsey SG;
 PI WPI; 2002-698627/75.
 DR N-PSDB; AAD45786.
 XX Treating and/or preventing neurodegenerative and neurodevelopmental
 PT disorders such as Alzheimer's, Parkinson's and Huntington's diseases by
 PT altering the ratio of amount of full-length and truncated TrkB or TrkC
 PT polypeptides.
 XX Claim 47; Page 56-59; 96pp; English.
 XX The present invention relates to a method of treating neurodegenerative
 CC or neurodevelopmental disorders in a mammal which involves administering
 CC an isolated nucleic acid encoding a full-length TrkB or TrkC or their
 CC mutant, variant, homologue or fragment or an anti-sense RNA for truncated
 CC TrkB or TrkC isoforms, where they increase the amount of full-length TrkB
 CC or TrkC or decrease the amount of truncated TrkB or TrkC in treated
 CC neurons. The methods and compositions of the invention are useful for
 CC treating or preventing neurodegenerative or neurodevelopmental disorders
 CC such as Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's
 CC disease (HD), amyotrophic lateral sclerosis (ALS), Lou Gehrig's disease),
 CC diabetic peripheral neuropathy, the adverse complications of Down's

CC syndrome (DS) and other types of peripheral neuropathy. Sequences of the
 CC invention are also used in gene therapy. The present sequence is human
 CC TrkB protein

XX SQ Sequence 822 AA;

Query Match 84.4%; Score 38; DB 5; Length 822;
 Best Local Similarity 80.0%; Pred. No. 6.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVRWH 5
 |::|||
 Db 4 WIRWH 8

RESULT 31

AAM50851
 ID AAM50851 standard; protein; 822 AA.

XX AC AAM50851;

XX DT 01-MAY-2002 (first entry)

XX DE Human receptor tyrosine kinase TrkB.

XX KW Receptor tyrosine kinase; TrkB; receptor; human;
 KW brain derived growth factor; BDNF; neurotrophin-3; NT-3;
 KW Huntington's disease; Parkinson's disease; Alzheimer's disease;
 KW amyotrophic lateral sclerosis; neurodegenerative disease; cancer;
 KW neuroprotective; neurotropic; anticonvulsant; antiparkinsonian; enzyme;
 KW cytosstatic; therapy.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

FT Peptide 1..31

FT Protein /label= Signal_peptide

FT Protein /label= Mature_protein

FT Domain 32..430

FT Modified-site /label= Extracellular

FT Modified-site 67

FT Domain /note= "N-glycosylated"

FT Region 72..117

FT Region /label= Leucine-rich_repeats

FT Modified-site 72..93

FT Modified-site /note= "leucine-rich repeat 1"

FT Region 95

FT Region /note= "N-glycosylated"

FT Modified-site 96..117

FT Modified-site /note= "leucine-rich repeat 2"

FT Modified-site 121

FT Modified-site /note= "N-glycosylated"

FT Modified-site 178

FT Modified-site /note= "N-glycosylated"

FT Domain 205

FT Domain /note= "N-glycosylated"

FT Domain 214..270

FT Modified-site /label= Ig-like_C2-type

FT Modified-site 241

FT Modified-site /note= "N-glycosylated"

FT Modified-site 254

FT Modified-site /note= "N-glycosylated"

FT Modified-site 280

FT Domain /note= "N-glycosylated"

FT Domain 301..365

FT Modified-site /label= Ig-like_C2-type

FT Modified-site 325

FT Modified-site /note= "N-glycosylated"

FT Modified-site 338

FT Modified-site /note= "N-glycosylated"

FT Modified-site 412

FT Modified-site /note= "N-glycosylated"

FT Domain 431..454
 FT /label= Transmembrane
 FT 455..822
 FT /label= Cytoplasmic
 FT Misc-difference 467..477
 FT /note= "replaced by FVLPHKIPLDG is truncated isoform"
 FT Misc-difference 478..822
 FT /note= "missing in truncated isoform"
 FT Binding-site 516
 FT /note= "interaction with SHC protein"
 FT Modified-site 516
 FT /note= "O-phosphorylated"
 FT Domain 538..807
 FT /label= Protein_kinase
 FT Region 544..552
 FT /note= "ATP binding site"
 FT Region 572
 FT /note= "ATP binding site"
 FT Active-site 676
 FT Modified-site 702
 FT /note= "O-phosphorylated"
 FT Modified-site 706
 FT /note= "O-phosphorylated"
 FT Modified-site 707
 FT /note= "O-phosphorylated"
 FT Binding-site 817
 FT /note= "interaction with PLC-gamma-1"
 FT Modified-site 817
 FT /note= "O-phosphorylated"
 PN WO200203071-A2.
 XX 10-JAN-2002.
 PD 05-JUL-2001; 2001WO-US021472.
 PF 05-JUL-2000; 2000US-0215778P.
 PR (PANG-) PANGENE CORP.
 PA Bates AT;
 XX WPI; 2002-179638/23.
 DR Screening for a neurotrophic factor mimetic, useful for treating, e.g.,
 PT cancer and Alzheimer's, comprises combining a candidate mimetic with a
 PT fragment of a tyrosine kinase protein.
 XX Disclosure; Fig 11A; 107pp; English.

The present sequence is that of human receptor tyrosine kinase TrkB, the
 receptor for brain-derived growth factor (BDNF) and neurotrophin-3 (NT-
 3). The invention concerns Trks and their ligands that modulate cell
 growth, differentiation and survival. Trk proteins are known to mediate
 the activities of neurotrophins and are also known proto-oncogenes.
 Methods are claimed for screening for small molecule neurotrophic factor
 (NTF) mimetics, such as the cyclic peptide given in AAM50844, capable of
 binding to a Trk protein or of modulating the binding of a neurotrophin
 to a Trk protein. Also claimed are medicaments comprising a small
 molecule NTF mimetic and their use in claimed methods for treatment of
 cancer or a neurodegenerative disease selected from Huntington's disease,
 Parkinson's disease, Alzheimer's disease and amyotrophic lateral
 sclerosis

XX SQ Sequence 822 AA;

Query Match 84.4%; Score 38; DB 5; Length 822;
 Best Local Similarity 80.0%; Pred. No. 6.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVRWH 5
 |::|||

Db 4 WIRWH 8

RESULT 32

ABU56698

ID ABU56698 standard; protein; 822 AA.

XX AC ABU56698;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polypeptide #291.

XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
XX KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX XX

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US012476.

XX PR 18-APR-2001; 2001US-0284770P.

XX PR 10-MAY-2001; 2001US-0290492P.

XX PR 09-NOV-2001; 2001US-0339245P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 29-NOV-2001; 2001US-0334370P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (EOSB-) BOS BIOTECHNOLOGY INC.

XX FI Aziz N, Murray R;

XX DR WPI; 2003-093161/08.

XX DR N-PSDB; ABX76427.

XX PT Detecting a lung cancer-associated transcript in a cell from a patient
XX PT for treating lung cancer, by contacting a biological sample from the
XX PT patient with a polynucleotide that exhibits increased or decreased
XX PT expression in lung cancer.

XX PS Claim 27; Page 417-418; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated
XX CC transcript in a cell from a patient, comprising contacting a biological
XX CC sample from the patient with a polynucleotide that selectively hybridizes
XX CC to a sequence that is at least 80 % identical to a gene that exhibits
XX CC increased or decreased expression in lung cancer samples. Lung cancer-
XX CC associated polynucleotides and polypeptides are used for identifying a
XX CC compound that modulates a lung cancer-associated polypeptide, for
XX CC inhibiting proliferation of a lung cancer-associated cell to treat lung
XX CC cancer in a patient and for treating a mammal having lung cancer by
XX CC administering a modulatory compound identified. The methods are useful
XX CC for treating lung cancer, such as small cell lung cancer, non-small cell
XX CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX CC for diagnostic purposes and as targets for screening for therapeutic
XX CC compounds that modulate lung cancer, such as antibodies. Sequences
XX CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
XX CC invention

XX SQ Sequence 822 AA;

Query Match 84.4%; Score 38; DB 6; Length 822;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy

1 WVRWH 5

|:|

4 WIRWH 8

Db

RESULT 33

ABR82949

ID ABR82949 standard; protein; 822 AA.

XX AC ABR82949;

XX DT 18-DEC-2003 (first entry)

XX DE Human TrkB polypeptide.

XX KW TrkB; TrkC; neurodegenerative; neuro-developmental; antiparkinsonian;
XX KW neuroprotective; anticonvulsant; cerebroprotective; vasotropic;
XX KW nootropic; human.

XX OS Homo sapiens.

XX XX

XX PN WO2003071872-A1.

XX PD 04-SEP-2003.

XX PF 28-MAY-2002; 2002WO-US016807.

XX PR 22-FEB-2002; 2002WO-US005151.

XX XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX PA (KRUE/) KRUEGER B K.

XX PA (KING/) KINGSBURY T J.

XX PA (BAMB/) BAMBRICK L L.

XX PA (DORS/) DORSEY S G.

XX PI Krueger BK, Kingsbury TJ, Bambrick LL, Dorsey SG;

XX DR WPI; 2003-731549/69.

XX DR N-PSDB; ACF36565.

XX PT Treating and/or preventing neurodegenerative or neuro-developmental
XX PT disorders, such as Alzheimer's disease, Parkinson's disease and
XX PT amyotrophic lateral sclerosis, using nucleic acids encoding TrkB and/or
XX PT TrkC polypeptides.

XX PS Claim 47; Page 58-61; 99pp; English.

XX CC The invention relates to treating a neurodegenerative or neuro-
XX CC developmental disorder in a mammal. The method involves altering the
XX CC ratio of the amount of full length TrkB polypeptide to the amount of
XX CC truncated TrkB polypeptides in a neuron or by altering the ratio of the
XX CC amount of full length TrkC polypeptide to the amount of truncated TrkC
XX CC polypeptides in a neuron. The methods and compositions of the present
XX CC invention are useful for treating and/or preventing a neurodegenerative
XX CC or neuro-developmental disorder, such as Alzheimer's disease, Parkinson's
XX CC disease, Huntington's disease, amyotrophic lateral sclerosis (Lou
XX CC Gehrig's disease), the adverse neurologic complications of Down syndrome,
XX CC diabetic peripheral neuropathy and other types of peripheral neuropathy,
XX CC and is associated with an injury to the central or peripheral nervous
XX CC system resulting from stroke, cerebral ischaemia, or chemical and/or
XX CC physical trauma. The present sequence represents a human TrkB polypeptide
XX CC (GenBank Accession No. NM_006180)

XX SQ Sequence 822 AA;

Query Match 84.4%; Score 38; DB 7; Length 822;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy

1 WVRWH 5

|:|

4 WIRWH 8

Db

RESULT 34
ADE40445
ID ADE40445 standard; protein; 822 AA.
XX
AC ADE40445;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human receptor tyrosine kinase TRKB (gene ID 5816) protein.
XX
KW AIDS; acquired immunodeficiency syndrome; human immunodeficiency virus;
KW HIV-related disorder; differential expression; drug screening;
KW viral replication modulation; diagnosis; prognosis; predisposition;
KW anti-HIV; gene therapy; antineoplastic therapy; human;
KW receptor tyrosine kinase TRKB; enzyme.
XX
OS Homo sapiens.
XX
FN WO2003070883-A2.
XX
PD 28-AUG-2003.
XX
PF 13-FEB-2003; 2003WO-US004246.
XX
PR 15-FEB-2002; 2002US-0357391P.
PR 13-MAY-2002; 2002US-0380249P.
PR 25-JUN-2002; 2002US-0391306P.
PR 27-AUG-2002; 2002US-0406297P.
PR 19-SEP-2002; 2002US-0412007P.
PR 10-OCT-2002; 2002US-0417508P.
PR 10-DEC-2002; 2002US-0432318P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Powell DM, Weich NS;
XX
WPI; 2003-671808/63.
DR N-PSDB; ADE40444.
XX
PT Identifying a compound capable of diagnosing, preventing or treating AIDS
PT or an HIV-related disorder comprises assaying the ability of the compound
PT to modulate e.g. 1414, 1481 or 1553 nucleic acid expression or
PT polypeptide activity.
XX
PS Claim 1; SEQ ID NO 24; 167pp; English.
XX
CC The invention relates to a method of identifying a compound useful in the
CC treatment of AIDS (acquired immunodeficiency syndrome) or an HIV (human
CC immunodeficiency virus)-related disorder. The invention involves assaying
CC the ability of a test compound to modulate the activity or expression of
CC 26 human proteins. These proteins and nucleic acids encoding them
CC (ADE40422-ADE40473) are differentially expressed in tissues relating to
CC AIDS or an HIV-related disorder compared to their expression in normal
CC tissues. The invention also relates to the use of the compounds
CC identified to modulate viral replication in a cell and to treat a patient
CC with AIDS or an HIV-related disorder. The invention further discloses
CC methods for the diagnostic evaluation and prognosis of various HIV-
CC related disorders, and for the identification of individuals exhibiting a
CC predisposition to such conditions. The modulatory compounds identified
CC using the method of the invention may be small organic molecules,
CC peptides, antibodies or antisense nucleic acid molecules. The methods of
CC the invention are useful in diagnosing, preventing or treating AIDS or
CC HIV-related disorders. The present sequence represents a human protein
CC which is differentially expressed in AIDS or HIV-related disorders.
XX
SQ Sequence 822 AA;
XX
Query Match 84.4%; Score 38; DB 7; Length 822;
Best Local Similarity 80.0%; Pred. NO. 6.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5

Db | : | | |
4 WIRWH 8
RESULT 35
ABU56699
ID ABU56699 standard; protein; 838 AA.
XX
AC ABU56699;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polypeptide #292.
XX
KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
FN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSE-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
WPI; 2003-093161/08.
DR N-PSDB; ABX76428.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 27; Page 418-419; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridizes
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC invention
XX
SQ Sequence 838 AA;
XX
Query Match 84.4%; Score 38; DB 6; Length 838;
Best Local Similarity 80.0%; Pred. NO. 6.4e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
|:|:|
Db 4 WIRWH 8

RESULT 36
AAB40277
ID AAB40277 standard; protein; 79 AA.
XX AC AAB40277;
XX DT 08-FEB-2001 (first entry)
XX DE Human ORFX ORF41 polypeptide sequence SEQ ID NO:82.
XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX OS Homo sapiens.
XX WO200058473-A2.
XX PN 05-OCT-2000.
XX PD 31-MAR-2000; 2000WO-US008621.
XX PF 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX WPI; 2000-602362/57.
XX N-PSDB; AAC74486.
XX PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
XX PS Claim 11; Page 465; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX dermatological; immunosuppressive; antinflammatory; antibacterial;
XX antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
XX sequences can be used for determining the presence of or predisposition
XX to, or preventing or treating pathological conditions associated with an
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX proteins in gene therapy vectors. The proteins and nucleic acids may be
XX used to treat cancers, proliferative disorders, neurodegenerative
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester

CC storage, systemic lupus erythematosus, severe combined immunodeficiency
(SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
cartilage damage, nocturnal haemoglobinuria, antinflammatory disease; to
enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX SQ Sequence 79 AA;

Query Match 80.0%; Score 36; DB 3; Length 79;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
|:|:|
Db 2 WLRWH 6

RESULT 37
ABP11225
ID ABP11225 standard; protein; 79 AA.
XX AC ABP11225;
XX DT 25-JUN-2002 (first entry)
XX DE Human ORFX protein sequence SEQ ID NO:22432.
XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX OS Homo sapiens.
XX WO200192523-A2.
XX PN 06-DEC-2001.
XX PF 29-MAY-2001; 2001WO-US010836.
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach MD;
XX WPI; 2002-106308/14.
XX N-PSDB; ABN26977.
XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
XX PS Disclosure; SEQ ID NO 22432; 1037pp; English.
XX
XX The present invention describes substantially purified human proteins
(referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
in the specification). ABN15762 to ABN27252 encode the human ORFX
proteins given in ABP0010 to ABP11500. ORFX proteins are useful for
treating or preventing a pathology associated with an ORFX-associated
disorder in humans, and in the manufacture of a medicament for treating a
syndrome associated with ORFX-associated disorder. ORFX polynucleotide
sequences can be used in gene therapy. ORFX sequences can be used in the
treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
osteoarthritis, neurodegenerative disorders, disorders related to organ
transplantation, cardiovascular diseases, diabetes mellitus, systemic
lupus erythematosus, hypertension, hypothyroidism, cholesterol ester

CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. OREFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. the sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 79 AA;
 SQ
 Query Match 80.0%; Score 36; DB 5; Length 79;
 Best Local Similarity 80.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db 2 WLRWH 6
 | : | | |
 | : | | |
 RESULT 38
 AAY25731
 ID AAY25731 standard; protein; 83 AA.
 AC AAY25731;
 XX
 DT 04-OCT-1999 (first entry)
 XX
 DE Human secreted protein encoded from gene 21.
 XX
 KW Secreted protein; human; treatment; diagnosis; therapy; cancer; tumour;
 KW neurodegenerative disorder; developmental abnormality; blood disorder;
 KW fetal deficiency; blood disorder; leukemia; immune system; inflammation;
 KW autoimmune disease; hepatic disease; renal disease; allergy; restenosis;
 KW ischaemic shock; Alzheimer's disease; cognitive disorder; schizophrenia;
 KW cardiovascular disorder; wound healing; stroke; arthritis; obesity;
 KW asthma; sepsis; acne; psoriasis; transplant rejection; infection; AIDS;
 KW metabolic disorder.
 XX
 OS Homo sapiens.
 XX
 FN WO9938881-A1.
 XX
 PD 05-AUG-1999.
 XX
 EF 27-JAN-1999; 99WO-US001621.
 XX
 PR 30-JAN-1998; 98US-0073159P.
 PR 30-JAN-1998; 98US-0073160P.
 PR 30-JAN-1998; 98US-0073161P.
 PR 30-JAN-1998; 98US-0073162P.
 PR 30-JAN-1998; 98US-0073164P.
 PR 30-JAN-1998; 98US-0073165P.
 PR 30-JAN-1998; 98US-0073167P.
 PR 30-JAN-1998; 98US-0073170P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Ferrie AM, Rosen CA, Florence KA, Carter KC, Soppet DR;
 PI Yu G, Florence C, Young P, Ni J, Feng P, Endress GA, Janat P;
 XX
 DR WPI; 1999-469315/39.
 DR N-PSDB; AAZ00430.
 XX
 PT New isolated human genes and the secreted polypeptides they encode useful
 PT in, e.g. treatment of Alzheimer's.
 XX
 PS Claim 1b; Page 309-310; 393pp; English.
 XX
 PS This invention describes novel human genes (see AAZ00410-Z00477) and the

CC secreted proteins (see AAY25711-Y25778) and fragments (see AAY25779-
 CC Y25907) they encode. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Specific uses are described for each of the 67
 CC polynucleotides of the invention, based on which tissues they are most
 CC highly expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumours, neurodegenerative disorders, developmental
 CC abnormalities and fetal deficiencies, blood disorders, leukemias,
 CC diseases of the immune system, autoimmune diseases, hepatic and renal
 CC diseases, inflammation, allergies, ischaemic shock, Alzheimer's and
 CC cognitive disorders, schizophrenia, restenosis, cardiovascular disorders,
 CC wound healing, stroke, arthritis, obesity, asthma, sepsis, acne,
 CC psoriasis, transplant rejection, metabolic disorders, infections and
 CC AIDS. The polypeptides are also useful for identifying their binding
 CC partners
 XX
 SQ Sequence 83 AA;
 Query Match 80.0%; Score 36; DB 2; Length 83;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WVRWHF 6
 Db 42 WVRWKF 47
 | | | | |
 | | | | |
 RESULT 39
 ABG07463
 ID ABG07463 standard; protein; 91 AA.
 AC ABG07463;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #7454.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 FN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS71650.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 37822; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed

genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 80.0%; Score 36; DB 4; Length 91;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVRWHF 6
Db 86 WVKWVF 91
|||:|

RESULT 40
AAE13431
ID AAE13431 standard; protein; 424 AA.
AC AAE13431;
DT 12-FEB-2002 (first entry)
DE Soybean microsomal omega-6 desaturase, gmFAD6 protein.
KW Soybean; delta9-desaturase; fatty acid delta12-desaturase; palmitic acid; linoleic acid; stearic acid; oleic acid; transgenic plant;
KW cotton seed oil; microsomal omega-6 desaturase; FAD6.
OS Glycine max.
FH Key Location/Qualifiers
FT Domain 165..170
FT /note= "Histidine binding motif"
FT Domain 201..205
FT /note= "Histidine binding motif"
FT Domain 361..365
FT /note= "Histidine binding motif"
XX WO200179499-A1.
XX
XX 25-OCT-2001.
XX
XX 18-APR-2001; 2001WO-AU000436.
XX
XX 18-APR-2000; 2000US-0198124P.
XX
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX Green A., Singh S, Liu Q;
XX WPI; 2001-602932/68.
XX
XX Modifying endogenous oil of cotton plants, to produce cotton seed oil with reduced palmitic and/or linoleic acid content, involves producing transgenic plants containing a fatty acid biosynthesis gene in a construct.
XX
XX Example 8; Fig 7; 201pp; English.
XX

CC The invention relates to a method for modifying the endogenous oil of a cotton plant, to produce cotton seed oil. The method comprises producing a transgenic cotton plant having a gene construct which includes a fatty acid biosynthesis gene operably linked to a promoter sequence capable of conferring expression of the delta9-desaturase (delta9 stearyl-ACP desaturase) gene, or fatty acid delta12-desaturase (oleoyl-PC delta12-desaturase) gene in the seed of a cotton plant. The invention is useful for producing cottonseed oil with reduced palmitic and/or linoleic acid content, and increased stearic and/or oleic acid content. The present sequence is soybean microsomal omega-6 desaturase, gmFAD6 protein related to the invention

SQ Sequence 424 AA;

Query Match 80.0%; Score 36; DB 4; Length 424;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WVRWHF 6
Db 251 WLMWHF 256
|:|

RESULT 41
ABB71663
ID ABB71663 standard; protein; 1169 AA.
AC ABB71663;
XX
XX 26-MAR-2002 (first entry)
DT
DE Drosophila melanogaster polypeptide SEQ ID NO 41781.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.

WO200171042-A2.
27-SEP-2001.
23-MAR-2001; 2001WO-US009231.
23-MAR-2000; 2000US-0191637P.
11-JUL-2000; 2000US-00614150.
(PEKE) PE CORP NY.
Venter JC, Adams M, Li PWD, Myers EW;
WPI; 2001-656860/75.
N-PSDB; ABL15766.
New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
Disclosure; SEQ ID NO 41781; 2lpp + Sequence Listing; English.
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB572072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Sequence 1169 AA;
SQ

Query Match 80.0%; Score 36; DB 4; Length 1169;
Best Local Similarity 80.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
Db 219 WLRWH 223

RESULT 42
AAE01508
ID AAB01508 standard; peptide; 6 AA.
XX
AC AAB01508;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Example; Page 26; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
SQ Sequence 6 AA;

Query Match 77.8%; Score 35; DB 3; Length 6;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 1 WVRWAF 6

RESULT 43
AAE21233
ID AAE21233 standard; protein; 53 AA.
XX
AC AAE21233;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human gene 18 encoded secreted protein HNNM45, SEQ ID NO:98.

XX Human, secreted protein; immune disorder; antiallergic; antirheumatic;
KW rheumatoid arthritis; breast neoplasia; breast cancer; antiarthritic;
KW neurological disease; Alzheimer's disease; Parkinson's disease; trauma;
KW Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania;
KW antiinflammatory; ophthalmological; dermatological; immunostimulatory;
KW immunomodulatory; immunosuppressive; antibacterial; antiporiatic;
KW gene therapy; autoimmune disease; Huntington's disease; meningitis;
KW demyelinating disease; peripheral neuropathy; congenital malformation;
KW spinal cord injury; peripheral neuropathy; ischaemia; perception;
KW multiple sclerosis; infarction; haemorrhage; schizophrenia; dementia;
KW depression; panic disorder; learning disability; ALS; feeding disorder;
KW hyperproliferative disorder; sleep pattern; cardiovascular disorder;
KW reproductive disorder; digestive system disorder; behavioural disorder.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT Protein /label= Signal_peptide
FT /label= Mature_secreted_protein
FT Misc-difference 52
FT /label= Unknown
FT /note= "Encoded by NCA"
XX
PN WO200216390-A1.
XX
PD 28-FEB-2002.
XX
PF 17-JAN-2001; 2001WO-US001435.
XX
PR 18-AUG-2000; 2000US-0226282P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Komatsoulis CA, Baker KP, Birse CE, Soppet DR;
PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
PI Fiscella M, Ni J;
XX
DR WPI; 2002-304113/34.
DR N-PSDB; AAD33734.
XX
PT An isolated nucleic acid molecule (I) comprising a polynucleotide which
PT encodes a polypeptide useful in the diagnosis and treatment of disorders
PT e.g. immune disorders.
PS Claim 11; Page 504; 504pp; English.
XX
CC AAD33692-AAD33736 represent cDNAs corresponding to 21 human secreted
CC protein genes, and AAE21191-AAE21235 represent the proteins they encode.
CC AAE21236-AAE21280 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 21 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of immune
CC or autoimmune diseases e.g. AIDS (acquired immune deficiency syndrome),
CC asthma, anaemia and rheumatoid arthritis, breast neoplasia and breast
CC cancer, neurological diseases e.g. Alzheimer's disease, Parkinson's
CC disease, Huntington's disease, Tourette syndrome, meningitis,
CC demyelinating disease, peripheral neuropathies, neoplasia, trauma,
CC congenital malformations, spinal cord injuries, toxic neuropathies
CC induced by neurotoxins, peripheral neuropathies, multiple sclerosis,
CC ischaemia and infarction, haemorrhages, schizophrenia, mania, dementia,
CC depression, panic disorder, learning disabilities, ALS, altered
CC behaviours e.g. disorders in feeding, sleep patterns, balance and
CC perception, encephalitis, disorders in cardiovascular, neural/ sensory,
CC reproductive and digestive systems, behavioural disorders and
CC hyperproliferative disorder. The present sequence represents a human
CC secreted protein of the invention


```

XX SQ Sequence 53 AA;
Query Match 77.8%; Score 35; DB 5; Length 53;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWHF 6
DB 21 WASWHF 26

RESULT 44
ABG64892
ID ABG64892 standard; protein; 53 AA.
XX
AC ABG64892;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human albumin fusion protein #1567.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfectivity; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US011988.
XX
PR 12-APR-2000; 2000US-0229358P.
XX
PR 25-APR-2000; 2000US-0199384P.
XX
PR 21-DEC-2000; 2000US-0256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
WPI; 2002-010886/01.
XX
New fusion protein for treating disease e.g. diabetes comprises an
albumin fused to a therapeutic protein.
XX
Claim 1; Page 1582; 2102pp; English.
XX
The present invention relates to albumin fusion proteins comprising a
therapeutic protein X and human albumin (HA, also known as human serum
albumin, HSA). The proteins are useful for treating a disease or disorder
that may be modulated by therapeutic protein X. The albumin extends the
shelf-life of protein X, and may increase its biological in vitro/in vivo
activity. The protein is useful for treating and diagnosing disorders
such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
disease, ulcerative colitis), immune disorders (e.g. acquired
immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
haematopoietic disorders, neural disorders (e.g. Alzheimer's,
Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
XX
ABG63326-ABG65518 represent albumin fusion proteins of the invention
XX
Sequence 53 AA;
Query Match 77.8%; Score 35; DB 5; Length 53;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;

XX SQ Sequence 53 AA;
Query Match 77.8%; Score 35; DB 4; Length 54;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 5
DB 1 WARWH 5

RESULT 45
AAO09345
ID AAO09345 standard; protein; 54 AA.
XX
AC AAO09345;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 23237.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
XX
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-514838/56.
XX
N-PSDB; AAI89276.
XX
Isolated nucleic acids and polypeptides, useful for preventing diagnosing
and treating e.g. leukemia, inflammation and immune disorders.
XX
Claim 20; SEQ ID NO 23237; 1399pp + Sequence Listing; English.
XX
The invention relates to human polynucleotides (AAI79941-AAI93841) and
the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activity/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
Sequence 54 AA;
Query Match 77.8%; Score 35; DB 4; Length 54;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 1 WARWH 5

RESULT 46
ABB11414
ID ABB11414 standard; peptide; 148 AA.

```

XX AC ABB11414;
 XX DT 11-JAN-2002 (first entry)
 XX DE Human breast cancer related protein BCRB2 homologue, SEQ ID NO:1784.
 XX KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder; disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 XX OS Homo sapiens.
 XX DN WO200157188-A2.
 XX PD 09-AUG-2001.
 XX PF 05-FEB-2001; 2001WO-US003800.
 XX PR 03-FEB-2000; 2000US-00496914.
 XX PR 27-APR-2000; 2000US-00560875.
 XX PA (HYSE-) HYSEQ INC.
 XX KW Tang YT, Liu C, Drmanac RT;
 DR WPI; 2001-457740/49.
 DR N-PSDB; ABA08658.
 XX PT Human proteins and DNA encoding sequences useful for preventing, treating
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.
 XX PS Claim 20; Page 184; 1963pp; English.
 XX CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic; thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,

CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness.
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention
 XX SQ Sequence 148 AA;

Query Match 77.8%; Score 35; DB 4; Length 148;
 Best Local Similarity 80.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVRWH 5
 Db 92 WVQWH 96
 ||:|
 ||:|

RESULT 47
 AAG83334
 ID AAG83334 standard; protein; 173 AA.

AC AAG83334;

DT 06-SEP-2001 (first entry)

DE P patens lipid metabolism related protein #52.

KW Moss; LMWP; lipid metabolism related protein; polyunsaturated fatty acid;
 KW fine chemical; transgenic plant.

OS Physcomitrella patens.

PN WO200138541-A1.

PD 31-MAY-2001.

PF 25-NOV-1999; 99WO-EP009108.

PR 25-NOV-1999; 99WO-EP009108.

PA (BADI) BASF PLANT SCI GMBH.

PI Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;
 DR WPI; 2001-381293/40.

PT New isolated nucleic acid molecule encoding lipid metabolism related
 PT proteins useful in the production of fine chemicals.

PS Claim 31; Page 105; 113pp; English.

XX The present invention provides the protein and coding sequences of a
 CC number of moss lipid metabolism related proteins (LMRPs). The moss
 CC Physcomitrella patens is one of the few plants able to produce
 CC polyunsaturated fatty acids, and the sequences can be used to create
 CC transgenic plants also capable of producing them. They can also be used
 CC to identify the presence of P. patens and in the production of fine
 CC chemicals. The present sequence is one of the proteins of the invention
 XX SQ Sequence 173 AA;

Query Match 77.8%; Score 35; DB 4; Length 173;
 Best Local Similarity 66.7%; Pred. No. 3.7e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WVRWH 6
 Db 73 WLLWH 78
 |:|
 |:|

```

RESULT 48
AAG80884
ID AAG80884 standard; protein; 173 AA.
XX
AC AAG80884;
XX
DT 28-AUG-2001 (first entry)
XX
DE Lipid modification protein protein sequence #5.
XX
KW Moss; Physcomitrella patens; lipid metabolism related protein; LMRP;
KW lipid biosynthesis; lipid modification; lipid degradation; cofactor;
KW fatty acid transport; genetic engineering; fatty acid; enzyme; plant;
KW microorganism: polynaturated fatty acid; oilseed plant; maize; wheat;
KW biotic stress tolerance; abiotic stress tolerance; rye; oat; triticale;
KW rice; barley; soybean; peanut; cotton; rapeseed; canola; manihot; pepper;
KW sunflower; tagetes; potato; tobacco; eggplant; tomato; Vicia; pea;
KW alfalfa; coffee; cacao; tea; Salix; oil palm; coconut; perennial grass;
KW forage crop.
XX
OS Physcomitrella patens.
XX
PN WO200138484-A2.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-EP011615.
XX
XX 25-NOV-1999; 99WO-EP009108.
XX
XX (BADI ) BASF PLANT SCI GMBH.
XX
XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;
XX Frank M, Freund A, Duwenig B, Schmidt R, Reski R;
XX WPI; 2001-367669/38.
XX
XX Nucleic acids encoding lipid metabolism related proteins from
XX Psychomitrella patens useful to produce fine chemicals in modified
XX organisms, particularly polynaturated fatty acids in oilseed plants.
XX
XX Claim 31; Page 114; 120pp; English.
XX
XX The present invention describes isolated nucleic acid sequences which
XX encode lipid metabolism related proteins (LMRP). The LMRP nucleic acids
XX can be used to modify lipids and fatty acids, cofactors and enzymes in
XX microorganisms and plants, particularly to produce polynaturated fatty
XX acids, and are especially useful in oilseed plants. The nucleic acids may
XX also confer biotic or abiotic stress tolerance, particularly to maize,
XX wheat, rye, oat, triticale, rice, barley, soybean, peanut, cotton,
XX rapeseed, canola, manihot, pepper, sunflower, tagetes, potato, tobacco,
XX eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao, tea, Salix
XX species, oil palm, coconut, perennial grasses and forage crops. AAH50878
XX to AAH50882 represent primers used in the exemplification of the present
XX invention. AAH50883 to AAH50968 represents LMRP nucleotide sequences, and
XX AAG80843 to AAG80928 represent LMRP protein sequences, given in the
XX present invention
XX
XX Sequence 173 AA;
XX
XX Query Match 77.8%; Score 35; DB 4; Length 173;
XX Best Local Similarity 66.7%; Pred. No. 3.7e+02;
XX Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 WVRWHF 6
XX | : |||
XX 73 WLLWHF 78
XX
XX RESULT 49
ADA54839
ADA54839 standard; protein; 190 AA.
ADA54839;
20-NOV-2003 (first entry)
Human protein, SEQ ID 2407.
Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
Gene Therapy; human; secretory protein; membrane proteins; cancer;
inflammatory disease; osteoporosis; neurological disease.
Homo sapiens.
EPI293569-A2.
19-MAR-2003.
21-MAR-2002; 2002EP-00006586.
14-SEP-2001; 2001JP-00328381.
24-JAN-2002; 2002US-0350435P.
(HELI-) HELIX RES INST.
(REAS-) RES ASSOC BIOTECHNOLOGY.
Isogai T, Sugiyama T, Otsuki T, Wakanatsu A, Sato H, Ishii S;
Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
Sexi N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
WPI; 2003-395539/38.
N-PSDB; ADA53200.
New polynucleotides encoding full-length polypeptides, e.g. secretory
and/or membrane proteins, useful for developing medicines for diseases in
which the gene is involved, or as target molecules for gene therapy.
Claim 14; SEQ ID NO 2407; 205pp; English.
The present invention relates to novel human secretory or membrane
proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
ADA54071). The coding sequences are useful in the gene therapy of
diseases caused by abnormalities of the proteins, e.g. cancer,
inflammatory diseases, osteoporosis or neurological disease.
Sequence 190 AA;
Query Match 77.8%; Score 35; DB 6; Length 190;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWHF 6
Dd 82 WGRWHF 87
RESULT 50
AAU51590
ID AAU51590 standard; protein; 663 AA.
AAU51590;
27-FEB-2002 (first entry)
Propionibacterium acnes immunogenic protein #12486.
SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
dermatological; osteopathic; neuroprotectant.
Propionibacterium acnes.

```

```

PN WO200181581-A2.
XX
PD
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59551.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Claim 3; SEQ ID NO 12785; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 663 AA;
SQ
Query Match 77.8%; Score 35; DB 4; Length 663;
Best Local Similarity 80.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db 289 WARWH 293
Search completed: June 10, 2004, 10:48:05
Job time : 63 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:43:44 ; Search time 11.6667 seconds
(without alignments)
49.470 Million cell updates/sec

Title: US-09-912-414-2
Perfect score: 45
Sequence: 1 WVRWHF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	88.9	443	T08136	probable omega-6 d
2	40	88.9	448	D85362	hypothetical prote
3	39	86.7	293	A83299	hypothetical prote
4	39	86.7	370	A96741	hypothetical prote
5	39	86.7	482	S49465	catalase (EC 1.11.
6	38	84.4	477	I73631	brain-derived neur
7	38	84.4	822	A56853	probable protein 5
8	37	82.2	591	G01586	hypothetical prote
9	36	80.0	185	T44538	probable permease
10	36	80.0	230	D70847	hypothetical prote
11	36	80.0	262	H83214	hypothetical prote
12	36	80.0	265	AC0441	probable aliphatic
13	36	80.0	266	A33238	hypothetical prote
14	36	80.0	384	E75295	conserved hypotet
15	36	80.0	424	T07742	omega-6 desaturase
16	36	80.0	472	AG2417	hypothetical prote
17	36	80.0	754	AE0614	hypothetical prote
18	36	80.0	1041	T29010	probable competenc
19	35	77.8	111	AB0690	hypothetical prote
20	35	77.8	226	F64450	probable membrane
21	35	77.8	458	B81409	hypothetical prote
22	35	77.8	508	G90372	probable transmemb
23	35	77.8	718	T29448	4-hydroxyphenylace
24	34	75.6	96	Q08C31	hypothetical prote
25	34	75.6	151	E64430	hypothetical prote
26	34	75.6	179	B83344	ribosomal protein
27	34	75.6	218	T27954	hypothetical prote
28	34	75.6	255	A69433	conserved hypotet
29	34	75.6	301	G83352	transcription regu

30	34	75.6	316	2	C47099	lyer homolog, nac
31	34	75.6	316	2	C84963	transcription regu
32	34	75.6	316	2	E85823	transcription regu
33	34	75.6	316	2	G90976	transcription regu
34	34	75.6	329	2	AF1891	hypothetical prote
35	34	75.6	358	1	S43876	site-specific DNA
36	34	75.6	358	2	A87296	modification methy
37	34	75.6	394	2	D90105	putative SAR DNA-b
38	34	75.6	424	2	JC5891	omega 6 desaturase
39	34	75.6	495	2	S52273	amine oxidase (fla
40	34	75.6	522	2	T40586	nucleolar protein
41	34	75.6	522	2	D96602	nucleolar protein
42	34	75.6	558	2	S75104	hypothetical prote
43	34	75.6	799	2	T48889	serine/threonine p
44	34	75.6	910	2	C69456	subtilisin sendai
45	33.5	74.4	1511	2	S60932	probable membrane
46	33	73.3	106	2	C82715	conserved hypotet
47	33	73.3	136	2	C95132	hypothetical prote
48	33	73.3	149	2	D84181	hypothetical prote
49	33	73.3	221	2	T27980	hypothetical prote
50	33	73.3	241	2	A71020	hypothetical prote
51	33	73.3	244	2	T15996	hypothetical prote
52	33	73.3	275	2	S76778	hypothetical prote
53	33	73.3	365	2	A85475	hypothetical prote
54	33	73.3	386	2	A97456	adenine-specific m
55	33	73.3	386	2	AD2674	adenine DNA methyl
56	33	73.3	396	1	E84987	bicyclomycin resis
57	33	73.3	396	2	B91013	bicyclomycin resis
58	33	73.3	396	2	D85857	bicyclomycin resis
59	33	73.3	396	2	AF0785	bicyclomycin resis
60	33	73.3	403	2	AF3432	site-specific DNA-
61	33	73.3	407	2	D86782	peptidoglycan bios
62	33	73.3	439	2	T01807	hypothetical prote
63	33	73.3	441	2	A83370	probable NFS trans
64	33	73.3	447	2	T17299	hypothetical prote
65	33	73.3	474	1	C39667	brain-derived neur
66	33	73.3	476	1	A35104	brain-derived neur
67	33	73.3	476	1	B39667	brain-derived neur
68	33	73.3	478	2	T09652	catalase (EC 1.11.
69	33	73.3	480	2	S76206	hypothetical prote
70	33	73.3	483	2	S37055	catalase (EC 1.11.
71	33	73.3	487	2	T32941	hypothetical prote
72	33	73.3	497	2	S27264	catalase (EC 1.11.
73	33	73.3	504	2	S48550	hypothetical prote
74	33	73.3	508	2	T50180	nucleolar protein
75	33	73.3	511	2	S58322	nucleolar protein
76	33	73.3	550	2	T06379	SAR DNA-binding pr
77	33	73.3	560	2	T06377	SAR DNA-binding pr
78	33	73.3	565	2	G82443	conserved hypotet
79	33	73.3	574	2	T29137	hypothetical prote
80	33	73.3	614	2	S84447	probable membrane
81	33	73.3	779	2	AG1978	hypothetical prote
82	33	73.3	821	1	S06943	brain-derived neur
83	33	73.3	821	1	A39667	brain-derived neur
84	33	73.3	1015	1	JG0628	formate dehydrogen
85	33	73.3	1015	2	C85729	hypothetical prote
86	33	73.3	1015	2	F90888	hypothetical prote
87	33	73.3	1016	1	S40838	formate dehydrogen
88	33	73.3	1016	2	D91231	formate dehydrogen
89	33	73.3	1016	2	AB0946	formate dehydrogen
90	33	73.3	1034	2	A95262	probable formate d
91	33	73.3	1164	2	T01871	RNA-directed DNA p
92	33	73.3	1607	2	T13250	hypothetical prote
93	33	73.3	2812	2	T43271	phosphotidylinosit
94	32.5	72.2	572	2	B83726	sulfite reductase
95	32	71.1	110	2	E81147	hypothetical prote
96	32	71.1	110	2	B81876	hypothetical prote
97	32	71.1	129	2	S48902	hypothetical prote
98	32	71.1	146	2	S84569	hypothetical prote
99	32	71.1	154	2	AB0163	probable membrane
100	32	71.1	210	2	D84016	hypothetical prote
101	32	71.1	243	2	S25755	ig lambda chain -
102	32	71.1	260	2	S76509	hypothetical prote

103	32	71.1	260	2	B71353	probable D, D-carbo	176	31	68.9	171	2	AD2457	hypothetical prote
104	32	71.1	283	2	T36769	hypothetical prote	177	31	68.9	206	2	T49138	hypothetical prote
105	32	71.1	287	2	G34758	yahB protein - Esc	178	31	68.9	226	2	C86255	protein F12P1.11
106	32	71.1	289	2	C83207	probable hydrolase	179	31	68.9	240	2	T37122	hypothetical prote
107	32	71.1	300	2	G96620	hypothetical prote	180	31	68.9	241	2	F72385	conserved hypothet
108	32	71.1	324	2	G69851	conserved hypothet	181	31	68.9	244	2	T01761	hypothetical prote
109	32	71.1	338	2	T49998	hypothetical prote	182	31	68.9	250	2	JE0157	mammalian sec22 pr
110	32	71.1	342	1	G70460	conserved hypothet	183	31	68.9	253	2	T01453	hypothetical prote
111	32	71.1	376	2	AG1420	hypothetical prote	184	31	68.9	270	2	C86295	hypothetical prote
112	32	71.1	376	2	AH1795	hypothetical membr	185	31	68.9	275	2	T43004	hypothetical prote
113	32	71.1	386	2	T12048	hypothetical membr	186	31	68.9	283	2	T14455	hypothetical prote
114	32	71.1	394	2	AC1554	DltB protein for D	187	31	68.9	286	2	AD1919	hypothetical prote
115	32	71.1	394	2	AE1196	DltB protein for D	188	31	68.9	288	2	T37081	hypothetical prote
116	32	71.1	395	2	S39659	DltB protein - Bac	189	31	68.9	289	2	AI2179	hypothetical prote
117	32	71.1	396	1	RXLL1A	ribosomal protein	190	31	68.9	295	1	S30383	hypothetical prote
118	32	71.1	396	1	RXLL1B	ribosomal protein	191	31	68.9	299	2	AC3087	morphine 6-dehydro
119	32	71.1	404	2	F89859	DltB membrane prot	192	31	68.9	304	2	T02125	transcription regu
120	32	71.1	407	2	T48308	60S ribosomal prot	193	31	68.9	305	2	F98199	chlorophyll a/b-bi
121	32	71.1	413	2	D95254	DltB protein (mpo	194	31	68.9	310	2	A25027	transcription regu
122	32	71.1	414	2	B98119	hypothetical prote	195	31	68.9	310	2	AB0275	arabinose operon r
123	32	71.1	421	1	JC4277	ribosomal protein	196	31	68.9	311	2	D86425	arabinose operon r
124	32	71.1	422	2	S24451	hypothetical prote	197	31	68.9	315	2	A95275	unknown protein [i
125	32	71.1	426	1	T09551	ribosomal protein	198	31	68.9	338	2	C70553	probable ABC trans
126	32	71.1	433	2	T48118	hypothetical prote	199	31	68.9	338	2	G71128	hypothetical prote
127	32	71.1	433	2	JH0594	hypothetical prote	200	31	68.9	341	2	S54079	hypothetical prote
128	32	71.1	459	2	JH0594	vasoactive intesti	201	31	68.9	352	2	H81165	snRNP-associated p
129	32	71.1	460	2	JC2194	vasoactive intesti	202	31	68.9	352	2	H87236	modification methy
130	32	71.1	473	2	H97092	closely related to	203	31	68.9	355	2	AI2867	endo-1,4-beta-xyla
131	32	71.1	474	2	F31283	catalase (EC 1.11.	204	31	68.9	358	2	H71088	hypothetical prote
132	32	71.1	477	2	AB2313	hypothetical prote	205	31	68.9	365	2	E97644	endo-1,4-beta-xyla
133	32	71.1	480	2	AB0148	catalase (EC 1.11.	206	31	68.9	368	2	AB2602	conserved hypothet
134	32	71.1	482	2	B83113	catalase PA4236 [i	207	31	68.9	368	2	B97384	probable integral
135	32	71.1	483	2	JH0532	catalase (EC 1.11.	208	31	68.9	377	2	E83420	hypothetical prote
136	32	71.1	484	2	A58663	catalase (EC 1.11.	209	31	68.9	383	2	C96581	hypothetical prote
137	32	71.1	495	2	JC2195	vasoactive intesti	210	31	68.9	408	2	S76830	phenoxylbenzoate di
138	32	71.1	495	2	D71308	probable alginat	211	31	68.9	409	2	S44171	L-rhamnose isomera
139	32	71.1	496	2	B83591	probable transport	212	31	68.9	418	2	H83843	PTS system galacti
140	32	71.1	499	2	A55227	probable transport	213	31	68.9	423	2	AH1407	PTS system galacti
141	32	71.1	503	2	AH2413	alginate O-acetyl	214	31	68.9	423	2	AH1783	probable En/Spn-11
142	32	71.1	503	2	E87101	probable membrane	215	31	68.9	428	2	F86459	hypothetical prote
143	32	71.1	503	2	D70930	hypothetical prote	216	31	68.9	430	2	AB2316	hypothetical prote
144	32	71.1	504	2	D81224	catalase (EC 1.11.	217	31	68.9	435	2	A70929	hypothetical prote
145	32	71.1	505	2	C64629	catalase (EC 1.11.	218	31	68.9	437	2	B87319	conserved phosphotr
146	32	71.1	505	2	F71885	catalase (EC 1.11.	219	31	68.9	452	2	H91172	probable Pts syste
147	32	71.1	507	2	I40767	catalase (EC 1.11.	220	31	68.9	452	2	AH0564	angio-associated m
148	32	71.1	507	2	G99908	Catalase [imported	221	31	68.9	452	2	I39383	cmat protein - Pse
149	32	71.1	507	2	AD3621	Catalase (EC 1.11.	222	31	68.9	453	2	B55543	conserved hypothet
150	32	71.1	508	2	T00753	probable cytochrom	223	31	68.9	457	2	G69341	glycosyltransferas
151	32	71.1	508	2	D84103	catalase (EC 1.11.	224	31	68.9	460	2	C82964	hypothetical prote
152	32	71.1	509	2	T05937	cytochrome P450 mo	225	31	68.9	462	2	H86018	hypothetical prote
153	32	71.1	520	2	A83203	alginate O-acetyl	226	31	68.9	465	2	S69038	hypothetical prote
154	32	71.1	527	2	G46426	alginate O-acetyla	227	31	68.9	475	2	B84171	probable permease
155	32	71.1	549	2	T02790	hypothetical prote	228	31	68.9	498	2	AD0854	cytochrome-c oxida
156	32	71.1	622	2	S61140	probable membrane	229	31	68.9	552	2	AC3435	gene NMB protein -
157	32	71.1	715	2	T26307	hypothetical prote	230	31	68.9	560	2	I38065	probable transposo
158	32	71.1	871	2	G84601	probable protein k	231	31	68.9	569	2	B85076	flagellar M-ring p
159	32	71.1	878	2	T08559	protein kinase hom	232	31	68.9	580	2	T82114	related to beta tr
160	32	71.1	886	2	T35469	probable ATP /GTP-	233	31	68.9	605	2	T49780	ferrous iron trans
161	32	71.1	1034	2	A24925	beta-galactosidase	234	31	68.9	614	1	S75294	probable spPA prot
162	32	71.1	1099	2	AE1065	conserved hypothet	235	31	68.9	623	2	G70644	hypothetical prote
163	32	71.1	1285	1	BTQPD	dermonecrotic toxi	236	31	68.9	624	2	T33868	hypothetical prote
164	32	71.1	1508	2	T27828	hypothetical prote	237	31	68.9	626	2	B72281	hypothetical prote
165	32	71.1	1519	2	T27829	hypothetical prote	238	31	68.9	640	2	T28631	Y4cd protein - Rhi
166	32	71.1	1535	2	S46224	peroxidasein - fru	239	31	68.9	640	2	T26820	hypothetical prote
167	31.5	70.0	956	2	A65072	hypothetical prote	240	31	68.9	653	2	H86373	protein T23823.16
168	31.5	70.0	956	2	B61098	probable dehydroge	241	31	68.9	657	2	T01301	RNA-directed DNA p
169	31	68.9	956	2	F85943	probable dehydroge	242	31	68.9	660	2	H70798	probable cation-tr
170	31	68.9	105	2	H62388	hypothetical prote	243	31	68.9	665	2	F84523	En/Spn-like transp
171	31	68.9	106	2	T44499	hypothetical prote	244	31	68.9	707	2	S77094	glycogen operon pr
172	31	68.9	116	2	AB3504	hypothetical prote	245	31	68.9	767	2	B85079	hypothetical prote
173	31	68.9	119	1	MXKRLB	sigma lbnS protein	246	31	68.9	780	2	H64830	probable membrane
174	31	68.9	119	2	B34829	sigma lbnS protein	247	31	68.9	780	2	B85617	hypothetical prote
175	31	68.9	169	2	F75253	conserved hypothet	248	31	68.9	780	2	D90753	hypothetical prote
			170	2	S64488	regulatory protein							

249	31	68.9	780	2	T29580	hypothenical prote	322	164	1	ZPECL	signal peptidase I
250	31	68.9	803	2	T10060	receptor-like prot	323	164	2	F90832	prolipo protein sig
251	31	68.9	825	2	S75173	hypothenical prote	324	164	2	E85483	signal peptidase I
252	31	68.9	828	2	G87584	hypothenical prote	325	166	2	AB0508	lipoprotein signal
253	31	68.9	840	2	H89911	oxacillin resistan	326	169	2	AI0058	signal peptidase I
254	31	68.9	856	1	VCLJFP	env polyprotein pr	327	196	2	T24815	hypothenical prote
255	31	68.9	887	2	E96491	hypothenical prote	328	210	2	B81343	hypothenical prote
256	31	68.9	901	2	E70778	probable aceE prot	329	216	2	A39926	hypothenical prote
257	31	68.9	907	2	T44715	probable pyruvate	330	216	2	A48343	hypothenical prote
258	31	68.9	926	2	T40679	hypothenical prote	331	216	2	AF0160	hypothenical p
259	31	68.9	936	2	B87115	pyruvate dehydrog	332	229	2	S73788	probable hydrolase
260	31	68.9	936	2	B87115	hypothenical prote	333	231	2	C69300	hypothenical prote
261	31	68.9	936	2	B87115	neural cell adhesi	334	232	2	G69080	conserved hypochet
262	31	68.9	936	2	B87115	probable SpoIIIE-f	335	241	2	A23266	hypothenical prote
263	31	68.9	936	2	B87115	probable membrane	336	246	2	S27727	hypothenical prote
264	31	68.9	936	2	B87115	DNA segregation AT	337	246	2	B86066	hypothenical prote
265	31	68.9	936	2	B87115	genome polyprotein	338	246	2	B86066	hypothenical prote
266	31	68.9	936	2	B87115	triacylglycerol li	339	246	2	H91219	hypothenical prote
267	31	68.9	936	2	B87115	hypothenical prote	340	246	2	G65183	hypothenical prote
268	31	68.9	936	2	B87115	conserved hypochet	341	246	2	AI0920	probable UDP-N-ace
269	31	68.9	936	2	B87115	hypothenical prote	342	257	1	S50294	probable UDP-N-ace
270	31	68.9	936	2	B87115	hypothenical prote	343	257	2	C83982	hypothenical prote
271	31	68.9	936	2	B87115	probable integral	344	259	2	S32432	hypothenical prote
272	31	68.9	936	2	B87115	conserved hypochet	345	261	2	C83157	toxigen cytB - Bacil
273	31	68.9	936	2	B87115	BH0982 hypochetica	346	261	2	S75492	6-phosphogluconola
274	31	68.9	936	2	B87115	hypothenical prote	347	266	2	AB1234	hypochet coat polysac
275	31	68.9	936	2	B87115	hypochet coat polysac	348	266	2	AB1234	hypochet coat polysac
276	31	68.9	936	2	B87115	succinate dehydrog	349	266	2	AB1234	hypochet coat polysac
277	31	68.9	936	2	B87115	succinate dehydrog	350	266	2	AB1234	hypochet coat polysac
278	31	68.9	936	2	B87115	succinate dehydrog	351	266	2	AB1234	hypochet coat polysac
279	31	68.9	936	2	B87115	succinate dehydrog	352	266	2	AB1234	hypochet coat polysac
280	31	68.9	936	2	B87115	succinate dehydrog	353	266	2	AB1234	hypochet coat polysac
281	31	68.9	936	2	B87115	succinate dehydrog	354	266	2	AB1234	hypochet coat polysac
282	31	68.9	936	2	B87115	succinate dehydrog	355	266	2	AB1234	hypochet coat polysac
283	31	68.9	936	2	B87115	succinate dehydrog	356	266	2	AB1234	hypochet coat polysac
284	31	68.9	936	2	B87115	succinate dehydrog	357	266	2	AB1234	hypochet coat polysac
285	31	68.9	936	2	B87115	succinate dehydrog	358	266	2	AB1234	hypochet coat polysac
286	31	68.9	936	2	B87115	succinate dehydrog	359	266	2	AB1234	hypochet coat polysac
287	31	68.9	936	2	B87115	succinate dehydrog	360	266	2	AB1234	hypochet coat polysac
288	31	68.9	936	2	B87115	succinate dehydrog	361	266	2	AB1234	hypochet coat polysac
289	31	68.9	936	2	B87115	succinate dehydrog	362	266	2	AB1234	hypochet coat polysac
290	31	68.9	936	2	B87115	succinate dehydrog	363	266	2	AB1234	hypochet coat polysac
291	31	68.9	936	2	B87115	succinate dehydrog	364	266	2	AB1234	hypochet coat polysac
292	31	68.9	936	2	B87115	succinate dehydrog	365	266	2	AB1234	hypochet coat polysac
293	31	68.9	936	2	B87115	succinate dehydrog	366	266	2	AB1234	hypochet coat polysac
294	31	68.9	936	2	B87115	succinate dehydrog	367	266	2	AB1234	hypochet coat polysac
295	31	68.9	936	2	B87115	succinate dehydrog	368	266	2	AB1234	hypochet coat polysac
296	31	68.9	936	2	B87115	succinate dehydrog	369	266	2	AB1234	hypochet coat polysac
297	31	68.9	936	2	B87115	succinate dehydrog	370	266	2	AB1234	hypochet coat polysac
298	31	68.9	936	2	B87115	succinate dehydrog	371	266	2	AB1234	hypochet coat polysac
299	31	68.9	936	2	B87115	succinate dehydrog	372	266	2	AB1234	hypochet coat polysac
300	31	68.9	936	2	B87115	succinate dehydrog	373	266	2	AB1234	hypochet coat polysac
301	31	68.9	936	2	B87115	succinate dehydrog	374	266	2	AB1234	hypochet coat polysac
302	31	68.9	936	2	B87115	succinate dehydrog	375	266	2	AB1234	hypochet coat polysac
303	31	68.9	936	2	B87115	succinate dehydrog	376	266	2	AB1234	hypochet coat polysac
304	31	68.9	936	2	B87115	succinate dehydrog	377	266	2	AB1234	hypochet coat polysac
305	31	68.9	936	2	B87115	succinate dehydrog	378	266	2	AB1234	hypochet coat polysac
306	31	68.9	936	2	B87115	succinate dehydrog	379	266	2	AB1234	hypochet coat polysac
307	31	68.9	936	2	B87115	succinate dehydrog	380	266	2	AB1234	hypochet coat polysac
308	31	68.9	936	2	B87115	succinate dehydrog	381	266	2	AB1234	hypochet coat polysac
309	31	68.9	936	2	B87115	succinate dehydrog	382	266	2	AB1234	hypochet coat polysac
310	31	68.9	936	2	B87115	succinate dehydrog	383	266	2	AB1234	hypochet coat polysac
311	31	68.9	936	2	B87115	succinate dehydrog	384	266	2	AB1234	hypochet coat polysac
312	31	68.9	936	2	B87115	succinate dehydrog	385	266	2	AB1234	hypochet coat polysac
313	31	68.9	936	2	B87115	succinate dehydrog	386	266	2	AB1234	hypochet coat polysac
314	31	68.9	936	2	B87115	succinate dehydrog	387	266	2	AB1234	hypochet coat polysac
315	31	68.9	936	2	B87115	succinate dehydrog	388	266	2	AB1234	hypochet coat polysac
316	31	68.9	936	2	B87115	succinate dehydrog	389	266	2	AB1234	hypochet coat polysac
317	31	68.9	936	2	B87115	succinate dehydrog	390	266	2	AB1234	hypochet coat polysac
318	31	68.9	936	2	B87115	succinate dehydrog	391	266	2	AB1234	hypochet coat polysac
319	31	68.9	936	2	B87115	succinate dehydrog	392	266	2	AB1234	hypochet coat polysac
320	31	68.9	936	2	B87115	succinate dehydrog	393	266	2	AB1234	hypochet coat polysac
321	31	68.9	936	2	B87115	succinate dehydrog	394	266	2	AB1234	hypochet coat polysac

genome polyprotein
polyprotein(NSI, N
alpha-2-macroglobu
alpha-2-macroglobu
sulfite reductase
hypothetical early
transposase homolo
hypothetical prote
conserved hypotet
hypothetical prote
NADH2 dehydrogenas
NADH2 dehydrogenas
probable membrane
P13 protein - Leuc
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypotet
hypothetical prote
hypothetical prote
hypothetical prote
probable chorismat
nxi protein - Arth
hypothetical prote
hypothetical prote
Clorf36 protein -
hypothetical prote
superoxide dismuta
probable transcrip
transcription regu

conserved hypotet
NADH2 dehydrogenas
hypothetical prote
sensor histidine k
hypothetical 51.5
probable transport
protein T17H7.3 [i
hypothetical prote
cog protein - Par
hypothetical prote
hypothetical prote
hypothetical prote
alpha-amylose [imp
ferredoxin, probab
conserved hypotet
catalase (EC 1.11.
probable chromatin
hypothetical prote
probable trpE prot
FixG protein - Rhi
FixG Iron sulfur m
methionine-tRNA li
fas-activated seri
pectinesterase (EC
hypothetical prote
conserved hypotet
probable ABC trans
probable membrane
v-type sodium ATP
FOG2 protein - Yea
probable membrane
hypothetical prote
serine/threonine-s
probable membrane
hypothetical prote
hypothetical prote
ferrioxamine B rec
hypothetical prote
probable outer mem
probable outer mem
DEAD box protein R
hypothetical prote
dipeptidyl-peptida
organic solvent to
phosphoenolpyruvat
H+-exporting ATPas
cation-transportin
outer membrane uen
hypothetical prote
probable gamma-ada
hypothetical prote
formate dehydrogen
isoleucyl-tRNA syn
hypothetical prote
pol polyprotein -
hypothetical prote
hypothetical prote
hypothetical prote
protein-tyrosine k
hypothetical prote
conserved hypotet
immunoglobulin-lik
hypothetical prote
DNA polymerase III
N conserved hypoth
coagulation factor
coagulation factor
coagulation factor
genome polyprotein

ALIGNMENTS

RESULT 1

T08136
probable omega-6 desaturase (EC 1.14.99.-) precursor, chloroplast - rape
C/Species: Brassica napus (rape)
C/Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 08-Oct-1999
C/Accession: T08136
R.Hitz, W.D.; Carlson, T.J.; Booth, J.R.; Kinney, A.J.; Stecca, K.L.; Yadav, N.S.
Plant Physiol. 105, 635-641, 1994
A/Title: Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA and its
A/Reference number: Z16109; MUID:94345008; PMID:8066133
A/Accession: T08136
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-443 <HIT>
A/Cross-references: EMBL:129214; NID:9457630; PIDN:AAA50157.1; PID:9457631
A/Experimental source: seed
C/Genetics:
A/Genome: nuclear
C/Superfamily: omega-3 fatty acid desaturase
C/Keywords: chloroplast; fatty acid metabolism; oxidoreductase

Query Match 88.9%; Score 40; DB 2; Length 443;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
|||
Db 252 WVRWHF 257

RESULT 2

D85362
hypothetical protein AT930950 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002

C:Accession: D85362
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; MUID:20083488; PMID:10617198
 A:Accession: D85362
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-448 <STO>
 A:Cross-references: GB:NC_001268; NID:g7269997; PIDN:CAB79813.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4G30950
 A:Map position: 4
 C:Superfamily: omega-3 fatty acid desaturase

Query Match 88.9%; Score 40; DB 2; Length 448;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 |||||
 Db 257 WVRWHF 262

RESULT 3
 A83299
 hypothetical protein PA2778 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 A:Accession: A83299
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, N.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: A83299
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-293 <STO>
 A:Cross-references: GB:AE004705; GB:AE004091; NID:g9948851; PIDN:AAG06166.1; GSPDB:GN00140
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2778

Query Match 86.7%; Score 39; DB 2; Length 293;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 |||||
 Db 101 WVRWHF 106

RESULT 4
 A96741
 hypothetical protein F14023.22 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 A:Accession: A96741
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chinn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizart, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A96741
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-370 <STO>
 A:Cross-references: GB:AE005173; NID:g7239511; PIDN:AAF43237.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F14023.22
 A:Map position: 1

Query Match 86.7%; Score 39; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 |||||
 Db 151 WVRWH 155

RESULT 5
 S49465
 catalase (EC 1.11.1.6) - nematode (Onchocerca volvulus)
 C:Species: Onchocerca volvulus
 C:Date: 01-Feb-1995 #sequence_revision 10-Nov-1995 #text_change 04-Mar-2000
 A:Accession: S49465
 R:Henkle-Duehrsen, K.J.
 submitted to the EMBL Data Library, October 1994
 A:Reference number: S49452
 A:Accession: S49465
 A:Molecule type: mRNA
 A:Residues: 1-482 <HEN>
 C:Superfamily: catalase
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
 F:55_94,128/Active site: His, Ser, Asn #status predicted
 F:338/Binding site: heme iron (Yyr) (axial ligand) #status predicted

Query Match 86.7%; Score 39; DB 2; Length 482;
 Best Local Similarity 66.7%; Pred. No. 48;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 |||||
 Db 211 WVRWHF 216

RESULT 6
 I73631
 brain-derived neurotrophic factor receptor precursor, short splice form - human
 N:Alternate names: neurotrophin receptor trkB; truncated receptor tyrosine kinase trkB
 C:Species: Homo sapiens (man)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 11-Jun-1999
 A:Accession: I73631; I38357; S66384; S44129
 R:Shelton, D.L.; Sutherland, J.; Gripp, J.; Camerato, T.; Armanini, M.P.; Phillips, H.S.; J. Neurosci. 15, 477-491, 1995
 A:Title: Human trks: molecular cloning, tissue distribution, and expression of extracellular reference number: I56557; MUID:95123473; PMID:7823156
 A:Accession: I73631
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-477 <RES>
 A:Cross-references: GB:S76474; NID:g913719; PIDN:AAB33110.1; PID:g913720
 R:Allen, S.J.; Dawbarn, D.; Eckford, S.D.; Wilcock, G.K.; Ashcroft, M.; Colebrook, S.M.; Neuroscience 60, 825-834, 1994
 A:Title: Cloning of a non-catalytic form of human trkB and distribution of messenger RNA
 A:Reference number: I38357; MUID:95021262; PMID:7936202
 A:Accession: I38357
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-477 <RES>
 A:Cross-references: EMBL:X75958; NID:g473007; PIDN:CAAS3571.1; PID:g473008
 R:Haniou, M.; Talvenheimo, J.; Le, J.; Katta, V.; Weicher, A.; Ronde, M.F. Arch. Biochem. Biophys. 322, 256-264, 1995
 A:Title: Extracellular domain of neurotrophin receptor trkB: disulfide structure, N-glycosylation
 A:Reference number: S66384; MUID:96004804; PMID:7574684
 A:Accession: S66384

A:Molecule type: protein
 A:Residues: 32-37,38-42,43-57,143-155,167-179,183-199,302-308,329-333,334-337,'X',339-344
 C:Comment: This form of the receptor is missing the protein kinase domain.
 C:Genetics:
 A:Gene: GDB:NTRK2; trkB
 A:Cross-references: GDB:127898; OMIM:600456
 A:Map position: 9q22.1-9q22.1
 C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycoprotein family; nerve growth factor receptor; glycoprotein; growth factor receptor; tandem repeat; t
 C:Keywords: alternative splicing; #status predicted <SIG>
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-477/Product: brain-derived neurotrophic factor receptor, short splice form #status
 F:32-435/Domain: extracellular #status predicted <EXT>
 F:67-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:436-452/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <LR
 F:436-477/Domain: transmembrane #status predicted <TMN>
 F:453-477/Domain: cytosolic #status predicted <CYT>
 F:32-38,36-45,152-176,154-194,218-266,302-345/Disulfide bonds: #status experimental
 F:67,95,121,178,205,241,254,280,325,338,350,412/Binding site: carbohydrate (Asn) (coval
 Query Match 84.4%; Score 38; DB 1; Length 477;
 Best Local Similarity 80.0%; Pred. No. 69;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 | : | | |
 Db 4 WIRWH 8
 | : | | |
 RESULT 7
 A56853
 brain-derived neurotrophic factor receptor precursor - human
 N:Alternate names: receptor tyrosine kinase trkB
 N:Contains: protein-tyrosine kinase (EC 2.7.1.12)
 C:Species: Homo sapiens (man)
 C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 11-Jun-1999
 C:Accession: A56853; I56557
 R:Nakagawara, A.; Liu, X.G.; Ikegaki, N.; White, P.S.; Yamashiro, D.J.; Nycum, L.M.; Bl
 Genomics 25, 538-546, 1995
 A:Title: Cloning and chromosomal localization of the human TRK-B tyrosine kinase recept
 A:Reference number: A56853; MUID:95309922; PMID:778998
 A:Accession: A56853
 A:Molecule type: mRNA
 A:Residues: 1-822 <NAK>
 A:Cross-references: GB:U12140; NID:9525313; PIDN:AA51371.1; PID:G530791
 R:Shelton, D.L.; Sutherland, J.; Gripp, J.; Camerato, T.; Armanini, M.P.; Phillips, H.S.
 J. Neurosci. 15, 477-491, 1995
 A:Title: Human trks: molecular cloning, tissue distribution, and expression of extracell
 A:Reference number: I56557; MUID:95123473; PMID:7823156
 A:Accession: I56557
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-822 <SHE>
 A:Cross-references: GB:S76473; NID:G913717; PIDN:AAB33109.1; PID:G913718
 C:Genetics:
 A:Gene: GDB:NTRK2; trkB
 A:Cross-references: GDB:127898; OMIM:600456
 A:Map position: 9q22.1-9q22.1
 C:Function:
 A:Description: regulation of nervous system development; receptor for brain-derived neur
 C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycop
 C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor
 inase
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-822/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>
 F:32-435/Domain: extracellular #status predicted <EXT>
 F:67-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:436-452/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <LR
 F:453-822/Domain: transmembrane #status predicted <TMN>
 F:453-822/Domain: cytosolic #status predicted <CYT>

F:536-814/Domain: protein kinase homology <KIN>
 F:544-552/Region: protein kinase ATP-binding motif
 F:67,95,121,178,205,241,254,280,325,338,350,412/Binding site: carbohydrate (Asn) (coval
 F:572/Active site: Lys #status predicted
 F:706/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicte
 F:817/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 84.4%; Score 38; DB 1; Length 822;
 Best Local Similarity 80.0%; Pred. No. 11e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 | : | | |
 Db 4 WIRWH 8
 | : | | |

RESULT 8

G01586

Probable protein 55.11 homolog p67 [similarity] - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 20-Jun-2000

C:Accession: G01586

R;Yew, F.

Submitted to the EMBL Data Library, December 1994

A:Reference number: G07854

A:Accession: G01586

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-591 <YEW>

A:Cross-references: EMBL:U18247; NID:G603224; PIDN:AA57339.1; PID:G603225

Query Match

82.2%; Score 37; DB 2; Length 591;

Best Local Similarity 66.7%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 6
 | : | | |Db 350 WCRWHY 355
 | : | | |

RESULT 9

T44538

hypothetical protein VR2 [imported] - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 06-Oct-2000

C:Accession: T44538; B83570

R;Kakayama, K.; Takashima, K.; Ighihara, H.; Shinomiya, T.; Kageyama, M.; Kanaya, S.; O;

submitted to the EMBL Data Library, August 1999

A:Description: Genetic relationship between bacteriocins and bacteriophages.

A:Reference number: Z22790

A:Accession: T44538

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-185 <NAK>

A:Cross-references: EMBL:AB030825; PIDN:BAA83153.1

A:Experimental source: strain PA01

R;Stover, C.K.; Phan, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A:Reference number: AB2950; MUID:20437337; PMID:10984043

A:Accession: B83570

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-185 <STO>

A:Cross-references: GB:AE004497; GB:AE004091; NID:G9946478; PIDN:AAG04005.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0616

C:Superfamily: Pseudomonas aeruginosa hypothetical protein VR2

Query Match

80.0%; Score 36; DB 2; Length 185;

Best Local Similarity 80.0%; Pred. No. 60;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
|:|:|
Db 42 WLRWH 46

RESULT 10
D70847
Hypothetical protein RV0059 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: D70847
R:Colé, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70847
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-230 <COL>
A:Cross-references: GB:AL021428; GB:AL123456; NID:G3261514; PIDN:CAAL6240.1; PID:G280871
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0059
C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV0059

Query Match 80.0%; Score 36; DB 2; Length 230;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 6
|:|:|
Db 26 WIVWH 31

RESULT 11
H83214
Probable permease of ABC transporter PA3443 [imported] - Pseudomonas aeruginosa (strain H83214)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83214
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83214
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <STO>
A:Cross-references: GB:AE004765; GB:AE004091; NID:G9949580; PIDN:AAG06831.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3443
C:Superfamily: Synchococcus nitrate transport protein nrtB

Query Match 80.0%; Score 36; DB 2; Length 262;
Best Local Similarity 80.0%; Pred. No. 82;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
|:|:|
Db 248 WLRWH 252

RESULT 12
AC0441
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: E75295
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Iversen, P.W.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

probable aliphatic sulfonates transport permease protein suuC [imported] - Yersinia pestis
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AC0441
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I.; Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <KUR>
A:Cross-references: GB:AU590842; PIDN:CAC93095.1; PID:G15981547; GSPDB:GN00175
C:Genetics:
A:Gene: ssuC
C:Superfamily: Synchococcus nitrate transport protein nrtB

Query Match 80.0%; Score 36; DB 2; Length 265;
Best Local Similarity 80.0%; Pred. No. 83;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
|:|:|
Db 250 WLRWH 254

RESULT 13
AD3238
Hypothetical protein nrtB [imported] - Agrobacterium tumefaciens (strain C58, Dupont) pl:
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AD3238
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, K.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.; ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AD3238
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <KUR>
A:Cross-references: GB:AB008690; PIDN:AAL46322.1; PID:G17744108; GSPDB:GN00189
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: nrtB
A:Gene: plasmid
C:Superfamily: Synchococcus nitrate transport protein nrtB

Query Match 80.0%; Score 36; DB 2; Length 266;
Best Local Similarity 80.0%; Pred. No. 83;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
|:|:|
Db 257 WLRWH 261

RESULT 14
E75295
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: E75295
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Iversen, P.W.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266
 A;Accession: B75295
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-384 <WHI>
 A;Cross-references: GB:AE002058; GB:AE000513; NID:G6460059; PIDN:AAF11801.1; PID:G646006
 A;Experimental source: strain R1
 C;Genetics:
 A;Gene: DR2253
 A;Map position: 1

Query Match 80.0%; Score 36; DB 2; Length 384;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 5
 | : |||
 Db 362 WLRWH 366

RESULT 15

T07742
 omega-6 desaturase, chloroplast - soybean
 C;Species: Glycine max [soybean]
 C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 08-Oct-1999
 C;Accession: T07742
 R;Hit: W.D.; Carlson, T.J.; Booth, J.R.; Kinney, A.J.; Stecca, K.L.; Yadav, N.S.
 Plant Physiol. 105, 635-641, 1994
 A;Title: Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA and its ex
 A;Reference number: Z16109; MUID:94345008; PMID:8066133
 A;Accession: T07742
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-424 <HT>
 A;Cross-references: EMBL:L29215; NID:G459961; PIDN:AAA50158.1; PID:G459962
 A;Experimental source: seed
 C;Genetics:
 A;Genome: nuclear
 C;Superfamily: omega-3 fatty acid desaturase
 C;Keywords: chloroplast

Query Match 80.0%; Score 36; DB 2; Length 424;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 | : |||
 Db 251 WLRWHF 256

RESULT 16

AG2417
 hypothetical protein all4895 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. Strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: AG2417
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AG2417
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-472 <KUR>
 A;Cross-references: GB:BA000019; PIDN:BAB76594.1; PID:G17134033; GSPDB:GN00179
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: all4895

Query Match 80.0%; Score 36; DB 2; Length 472;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 | : |||
 Db 285 WYQWHF 290

RESULT 17

AE0614
 probable competence-related protein STY0984 [imported] - Salmonella enterica subsp. ente
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 25-Aug-2003
 C;Accession: AE0614
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AE0614
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-754 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD05383.1; PID:G16502146; GSPDB:GN00176
 C;Genetics:
 A;Gene: STY0984
 C;Superfamily: competence protein ComEC

Query Match 80.0%; Score 36; DB 2; Length 754;
 Best Local Similarity 80.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 5
 | : |||
 Db 542 WLRWH 546

RESULT 18

T29010
 hypothetical protein ZK328.1 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T29010
 R;Favell, A.

submitted to the EMBL Data Library, March 1996
 A;Description: The sequence of C. elegans cosmid ZK328.
 A;Reference number: Z20552

A;Accession: T29010
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1041 <FAV>
 A;Cross-references: EMBL:U50193; PIDN:AAA91251.1; CESP:ZK328.1
 A;Experimental source: strain Bristol N2
 C;Genetics:

A;Gene: CESP:ZK328.1

A;Introns: 25/1; 64/1; 143/1; 231/2; 285/3; 487/1; 567/2; 643/1; 937/3; 987/3; 1018/3

Query Match 80.0%; Score 36; DB 2; Length 1041;
 Best Local Similarity 80.0%; Pred. No. 2.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 5
 | : |||
 Db 316 WLRWH 320

RESULT 19

AB0690
 probable membrane protein STY1646 [imported] - Salmonella enterica subsp. enterica serov
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C/Accession: AB0690
 R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J. S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A/Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Agona
 A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A/Accession: AB0690
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-111 <PAR>
 A/Cross-references: GB:AL513382; PIDN:CAD01891.1; PID:g16502735; GSPDB:GNO0176
 C/Genetics:
 A/Gene: STY1646
 C/Superfamily: Escherichia coli probable membrane protein ydgC

Query Match 77.8%; Score 35; DB 2; Length 111;
 Best Local Similarity 60.0%; Pred. No. 54;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 | : : :
 Db 107 WIKWH 111

RESULT 20
 F64450
 Hypothetical protein MJ1207 - Methanococcus jannaschii
 C/Species: Methanococcus jannaschii
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 A/Accession: F64450
 R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A/Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
 A/Reference number: A64300; MUID:96337999; PMID:8688087
 A/Accession: F64450
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-226 <BUL>
 A/Cross-references: GB:U67562; GB:L77117; NID:g2826374; PIDN:AAB99211.1; PID:g1591837; T
 C/Genetics:
 A/Map position: FOR1150579-1151259
 C/Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1207

Query Match 77.8%; Score 35; DB 1; Length 226;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 | : : :
 Db 109 WARWF 114

RESULT 21
 B81409
 Probable transmembrane transport protein Cj0611c [imported] - Campylobacter jejuni (strain B81409)
 C/Species: Campylobacter jejuni
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C/Accession: B81409
 R/Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell, D.G.
 Nature 403, 665-668, 2000
 A/Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypoxanthine auxotrophy
 A/Reference number: AB1250; MUID:20150912; PMID:10688204
 A/Accession: B81409
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-458 <PAR>

A/Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CAB75247.1; PID:g696807
 A/Experimental source: serotype O2, strain NCTC 11168
 C/Genetics:
 A/Gene: Cj0611c

Query Match 77.8%; Score 35; DB 2; Length 458;
 Best Local Similarity 60.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
 | : : :
 Db 271 WARWH 275

RESULT 22
 G90372
 4-hydroxyphenylacetate-3-hydroxylase (hpaA) [imported] - *Sulfolobus solfataricus*
 C/Species: *Sulfolobus solfataricus*
 C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 01-Mar-2002
 C/Accession: G90372
 R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-arrett, R.A.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
 submitted to GenBank, April 2001
 A/Description: *Sulfolobus solfataricus* complete genome.
 A/Reference number: A99139
 A/Accession: G90372
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-508 <KUR>
 A/Cross-references: GB:AB006641; NID:gl3815338; PIDN:AAK42238.1; GSPDB:GNO0155
 C/Genetics:
 A/Gene: hpaA
 C/Superfamily: Escherichia coli 4-hydroxyphenylacetate 3-monooxygenase large chain

Query Match 77.8%; Score 35; DB 2; Length 508;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 | : : :
 Db 310 WFNWHF 315

RESULT 23
 T29448
 Hypothetical protein F08F3.2 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 12-Jun-2003
 C/Accession: T29448
 R/Blanchard, M.; Bradshaw, H.
 submitted to the EMBL Data Library, July 1996
 A/Description: The sequence of *C. elegans* cosmid F08F3.
 A/Reference number: Z20620
 A/Accession: T29448
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-718 <BLA>
 A/Cross-references: EMBL:U64847; PIDN:AAB04876.1; GSPDB:GNO0023; CESP:F08F3.2
 A/Experimental source: strain Bristol N2; clone F08F3
 C/Genetics:
 A/Gene: CESP:F08F3.2
 A/Map position: 5
 A/Introns: 42/3; 65/2; 156/3; 279/1; 310/3; 346/3; 406/1; 516/1; 552/3; 600/3; 667/1
 C/Superfamily: glycerol-3-phosphate O-acyltransferase

Query Match 77.8%; Score 35; DB 2; Length 718;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 | : : :
 Db 178 WCNWHF 183

```

RESULT 24
QOEC31
hypothetical protein C-96 rspA 3'-region - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 29-Sep-1999 #sequence_revision 29-Sep-1999 #text_change 29-Sep-1999
C;Accession: A04448
R;Schmieder, J.; Isono, K.
Nucleic Acids Res. 10, 1857-1865, 1982
A;Title: The DNA sequence of the gene rspA of Escherichia coli coding for ribosomal protein C
A;Reference number: A04448; MUID:82196866; PMID:6281725
A;Accession: A04448
A;Molecule type: DNA
A;Residues: 1-96 <SCH>
A;Cross-references: GB:V00352; GB:J01681; GB:J01682; NID:942899
C;Comment: This is the hypothetical translation of a sequence that was not reported as a
C;Genetics:
A;Map position: 20 min

Query Match 75.6%; Score 34; DB 4; Length 96;
Best Local Similarity 80.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
| | | |
Db 18 WVRWH 22

RESULT 25
B64430
ribosomal protein S12 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 12-Jun-2003
C;Accession: B64430
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: B64430
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-151 <BUL>
A;Cross-references: GB:U67547; GB:L77117; NID:91591695; PID:g1591700; TIGR:MJ1046; PID:g
C;Genetics:
A;Map position: FOR982651-983106
A;Start codon: GTG
C;Superfamily: ribosomal protein S12
C;Keywords: protein biosynthesis; ribosome
F;122/Modified site: beta-methylthiospartic acid (Asp) #status predicted

Query Match 75.6%; Score 34; DB 2; Length 151;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
| | | |
Db 25 WVRWH 29

RESULT 26
B83344
hypothetical protein PA2404 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: B83344
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bn
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: B83344
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-179 <STO>
A;Cross-references: GB:AE004668; GB:AE004091; NID:g9948446; PIDN:AA05792.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2404

Query Match 75.6%; Score 34; DB 2; Length 179;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 6
| | | |
Db 17 WVRWH 22

RESULT 27
T27954
hypothetical protein ZK669.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27954
R;Thomas, K.
submitted to the EMBL Data Library, September 1994
A;Reference number: Z20446
A;Accession: T27954
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-218 <WIL>
A;Cross-references: EMBL:Z37093; PIDN:CAA85464.1; GSPDB:GN00020; CESP:ZK669.3
C;Genetics:
A;Gene: CESP:ZK669.3
A;Map position: 2
A;Introns: 46/2; 88/2; 121/1; 148/2; 178/3

Query Match 75.6%; Score 34; DB 2; Length 218;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 6
| | | |
Db 47 WVRWH 52

RESULT 28
A69433
conserved hypothetical protein AFL466 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: A69433
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: A69433
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-255 <KLE>
A;Cross-references: GB:AE001001; GB:AE000782; NID:g2689324; PIDN:AA89780.1; PID:g26491

Query Match 75.6%; Score 34; DB 2; Length 255;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 WVRWH 5
|
|
|
|
Db 84 WFRWH 88

RESULT 29

G83352
transcription regulator MtlR PA2337 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: G83352
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83352
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-301 <STO>
A;Cross-references: GB:AE004660; GB:AE004091; NID:g9948372; PIDN:AAG05725.1; GSPDB:GN00145
A;Experimental source: strain PA01
C;Genetics:
A;Gene: mtlR; PA2337

Query Match 75.6%; Score 34; DB 2; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRWHF 6
|
|
|
|
Db 35 VRWHF 39

RESULT 30

C47099
lyeR homolog, nac 3'-region - Klebsiella pneumoniae
C;Species: Klebsiella pneumoniae
C;Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 26-Aug-1999
C;Accession: C47099
R;Schwacha, A.; Bender, R.A.
J. Bacteriol. 175, 2107-2115, 1993
A;Title: The nac (nitrogen assimilation control) gene from Klebsiella aerogenes.
A;Reference number: A47099; MUID:93209957; PMID:8458853
A;Accession: C47099
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-316 <SCH>
A;Cross-references: GB:L01114; NID:g149241; PIDN:AAA18174.1; PID:g149244
C;Genetics:
A;Start codon: GTG
C;Superfamily: regulatory protein lyeR
C;Keywords: DNA binding; transcription factor

Query Match 75.6%; Score 34; DB 2; Length 316;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
|
|
|
|
Db 163 WFRWH 167

RESULT 31

C64963
transcription regulator cbl - Escherichia coli (strain K-12)
N;Alternate names: cysB protein homolog
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: C64963; I41150
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C64963

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

A;Residues: 1-316 <BLAT>

A;Cross-references: GB:AE000290; GB:U00096; NID:g2367125; PIDN:AAC75049.1; PID:g1788296;
A;Experimental source: strain K-12, substrain MG1655

R;Komine, Y.; Adachi, T.; Inokuchi, H.; Ozeki, H.

J. Mol. Biol. 212, 579-598, 1990

A;Title: Genomic organization and physical mapping of the transfer RNA genes in Escherichia coli K-12
A;Reference number: I41149; MUID:90230300; PMID:2184240

A;Accession: I41150

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-142, 'R', 144-316 <RES>

A;Cross-references: GB:L31639; NID:g1004098; PIDN:AAA99929.1; PID:g1004100

C;Genetics:

C;Gene: cbl

C;Superfamily: regulatory protein lyeR

C;Keywords: DNA binding; transcription regulation

Query Match 75.6%; Score 34; DB 2; Length 316;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
|
|
|
|
Db 163 WFRWH 167

RESULT 32

E85823
transcription regulator cys regulon [imported] - Escherichia coli (strain O157:H7, subsp. C)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C;Accession: E85823
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85823
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-316 <STO>
A;Cross-references: GB:AE005174; NID:g12516162; PIDN:AAG57049.1; GSPDB:GN00145; UWGP:Z31
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: cbl
C;Superfamily: regulatory protein lyeR

Query Match 75.6%; Score 34; DB 2; Length 316;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
|
|
|
|
Db 163 WFRWH 167

RESULT 33

G90976
transcription regulator cys regulon [imported] - Escherichia coli (strain O157:H7, subsp. C)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
C;Accession: G90976
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gisawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic
A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: G90976
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-316 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA036206.1; PID:gl3362251; GSPDB:GN00154
A;Experimental source: strain 0157:H7, substrain RMD 050952
C;Genetics:
A;Gene: ECs2783
C;Superfamily: regulatory protein lysr

Query Match 75.6%; Score 34; DB 2; Length 316;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
| | | |
Db 163 WFRWH 167

RESULT 34
AF1891
hypothetical protein alr0679 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AF1891
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF1891
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-329 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA072637.1; PID:gl7130025; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0679

Query Match 75.6%; Score 34; DB 2; Length 329;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
| : | |
Db 225 WIQWH 229

RESULT 35
S43876
site-specific DNA methyltransferase (EC 2.1.1.1) - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 22-Jun-2003
C;Accession: S43876
R;Zweiger, G.; Marczynski, G.; Shapiro, L.
J. Mol. Biol. 235, 472-485, 1994
A;Title: A Caulobacter DNA methyltransferase that functions only in the predivisional cell
A;Reference number: S43876; MUID:94118303; PMID:8289276
A;Accession: S43876
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-358 <ZWE>
A;Cross-references: EMBL:U01032; NID:G939011; PIDN:AAAI8913.1; PID:G939012
C;Superfamily: type II site-specific DNA-methyltransferase
C;Keywords: methyltransferase; restriction modification system

Query Match 75.6%; Score 34; DB 1; Length 358;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWH 6
| | | |
Db 169 WYSWH 174

Db 332 WTYWHF 337

RESULT 36
A87296
modification methylase CcrMI [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 22-Jun-2003
C;Accession: A87296
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87296
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-358 <STO>
A;Cross-references: GB:AE005673; NID:gl3421535; PIDN:AAK22365.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC0378
C;Superfamily: type II site-specific DNA-methyltransferase

Query Match 75.6%; Score 34; DB 2; Length 358;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWH 6
| | | |
Db 332 WTYWHF 337

RESULT 37
D90105
putative SAR DNA-binding protein-1 [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: D90105
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re:
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: D90105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <DOU>
A;Cross-references: GB:AJ010592; NID:gl2580672; PIDN:CAC26989.1; GSPDB:GN00151
C;Genetics:
A;Map position: 2
A;Genome: nucleomorph
C;Keywords: nucleomorph

Query Match 75.6%; Score 34; DB 2; Length 394;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWH 6
| | | |
Db 169 WYSWH 174

RESULT 38
JC5891
omega 6 desaturase (EC 1.14.99.-) precursor, chloroplast - Chlamydomonas reinhardtii
C;Species: Chlamydomonas reinhardtii
C;Date: 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 20-Jun-2000
C;Accession: JC5891
R;Sato, N.; Fujiwara, S.; Kawaguchi, A.; Tsuzuki, M.
J. Biochem. 122, 1224-1232, 1997
A;Title: Cloning of a gene for chloroplast omega 6 desaturase of a green alga, Chlamydo
A;Reference number: JC5891; MUID:98158334; PMID:9498569

A:Accession: JC5891
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-424 <SAT>
A:Cross-references: DDBJ:AB007640; NID:g2696716; PIDN:BAA23881.1; PID:g2696717
C:Comment: This enzyme catalyzes the desaturation of monoenoic to dienoic acids.
C:Genetics:
A:Gene: des6
A:Superfamily: omega-3 fatty acid desaturase
C:Keywords: chloroplast; oxidoreductase
F:1-40/Domain: transit peptide (chloroplast) #status predicted <TPS>
Query Match 75.6%; Score 34; DB 2; Length 424;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWHF 6
DB 234 WLWVHF 239
RESULT 39
S55273
N:Alternate names: monoamine oxidase N
C:Species: Aspergillus niger
C:Date: 14-Oct-1995 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000
C:Accession: S55273; S55267; S60356
R:Schilling, B.; Lerch, K.
Mol. Gen. Genet. 247, 430-438, 1995
A:Title: Cloning, sequencing and heterologous expression of the monoamine oxidase gene f
A:Reference number: S55267; MUID:95287865; PMID:7770050
A:Accession: S55273
A:Molecule type: DNA
A:Residues: 1-495 <SCH>
A:Cross-references: EMBL:L38858; NID:g619754; PIDN:AAA98490.1; PID:g619755
A:Accession: S55267
A:Molecule type: protein
A:Residues: 165-175 <SCW>
R:Schilling, B.; Lerch, K.
Biochim. Biophys. Acta 1243, 529-537, 1995
A:Title: Amine oxidases from Aspergillus niger: identification of a novel flavin-depende
A:Reference number: S60356; MUID:95244610; PMID:7727530
A:Accession: S60356
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
C:Genetics:
A:Introns: 62/1; 89/3
C:Keywords: PAD; flavoprotein; oxidoreductase; peroxisome
F:41-69/Region: beta-alpha-beta FAD nucleotide-binding fold
F:493-495/Region: peroxisome/glyoxysome location signal #status atypical
Query Match 75.6%; Score 34; DB 2; Length 495;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWH 5
DB 94 WVWHH 98
RESULT 40
T40586
nucleolar protein involved in pre-rRNA processing - fission yeast (Schizosaccharomyces p
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T40586
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
Submitted to the EMBL Data Library, January 1999
A:Reference number: Z21938
A:Accession: T40586
A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-497 <SSEE>
A:Cross-references: EMBL:AL035216; PIDN:CAA22814.1; GSPDB:GN00067; SPDB:SPBC646.10c
A:Experimental source: strain 972h-; cosmid c646
C:Genetics:
A:Gene: SPDB:SPBC646.10c
A:Map position: 2
C:Superfamily: garden pea SAR DNA-binding protein
Query Match 75.6%; Score 34; DB 2; Length 497;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WVRWHF 6
DB 194 WYSWHF 199
RESULT 41
D96602
nucleolar protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D96602
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96602
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <STO>
A:Cross-references: GB:AE005173; NID:g6056371; PIDN:AAF02835.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: garden pea SAR DNA-binding protein
Query Match 75.6%; Score 34; DB 2; Length 522;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WVRWHF 6
DB 195 WYSWHF 200
RESULT 42
S75104
hypothetical protein sl10236 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S75104
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <KAN>
A:Cross-references: EMBL:D90910; GB:AB001339; NID:g1652956; PIDN:BAAL7966.1; PID:dl01863
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 75.6%; Score 34; DB 2; Length 558;
 Best Local Similarity 80.0%; Pred. No. 3.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VWRWH 5
 |||||
 Db 359 VWRWH 363

RESULT 43
 T48889
 serine/threonine protein kinase (EC 2.7.1.1-) afSK [validated] - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 20-Jan-2003
 C;Accession: T48889
 R;Matsumoto, A.; Hong, S.; Ishizuka, H.; Horinouchi, S.; Beppu, T.
 Gene 146, 47-56, 1994
 A;Title: Phosphorylation of the AfSR protein involved in secondary metabolism in Streptococcus
 A;Reference number: Z24852; MUID:94341568; PMID:8063104
 A;Accession: T48889
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-799 <MAT>
 A;Cross-references: EMBL:D15062; PIDN:BAB40935.1
 A;Experimental source: strain M130
 C;Genetics:
 A;Gene: afSK
 C;Function:
 A;Description: EC 2.7.1.1- [validated, MUID:94341568]; phosphorylates AfSR, a global regulator
 C;Keywords: phosphotransferase

Query Match 75.6%; Score 34; DB 2; Length 799;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRWHF 6
 |||||
 Db 628 VRWHF 632

RESULT 44
 C69456
 subtilisin sendai homolog - Archaeoglobus fulgidus
 C;Species: Archaeoglobus fulgidus
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
 C;Accession: C69456
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
 A;Reference number: A69250; MUID:98049343; PMID:9389475
 A;Accession: C69456
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-910 <KLE>
 A;Cross-references: GB:AE000989; GB:AE000782; NID:92689312; PIDN:AAB89591.1; PID:9264889

Query Match 75.6%; Score 34; DB 2; Length 910;
 Best Local Similarity 66.7%; Pred. No. 5.3e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VWRWHF 6
 |||||
 Db 408 VWRWHF 413

RESULT 45
 S60932
 probable membrane protein YPL058c - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein LPE14c
 C;Species: Saccharomyces cerevisiae
 C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 02-Feb-2001
 C;Accession: S60932
 R;Winnett, E.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.;
 submitted to the EMBL Data Library, October 1995
 A;Reference number: S60921
 A;Accession: S60932
 A;Molecule type: DNA
 A;Residues: 1-1511 <WIN>
 A;Cross-references: EMBL:U39205; NID:91079672; PIDN:AAB68307.1; PID:91079684; MIPS:YPL05
 C;Genetics:
 A;Gene: SGD:PDR12
 A;Cross-references: SGD:S0005979; MIPS:YPL058c
 A;Map position: 16L
 C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C;Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
 F;154-173/Domain: ATP-binding cassette homology <ABC1>
 F;513-529/Domain: transmembrane #status predicted <TM1>
 F;549-565/Domain: transmembrane #status predicted <TM2>
 F;602-618/Domain: transmembrane #status predicted <TM3>
 F;659-675/Domain: transmembrane #status predicted <TM4>
 F;766-782/Domain: transmembrane #status predicted <TM5>
 F;861-1060/Domain: ATP-binding cassette homology <ABC2>
 F;878-885/Region: nucleotide-binding motif A (P-loop)
 F;991-1007/Domain: transmembrane #status predicted <TM6>
 F;1184-1200/Domain: transmembrane #status predicted <TM7>
 F;1208-1224/Domain: transmembrane #status predicted <TM8>
 F;1290-1306/Domain: transmembrane #status predicted <TM9>
 F;1324-1340/Domain: transmembrane #status predicted <TM10>
 F;1450-1466/Domain: transmembrane #status predicted <TM11>

Query Match 74.4%; Score 33.5; DB 2; Length 1511;
 Best Local Similarity 71.4%; Pred. No. 1e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 VWRWHF 6
 |||||
 Db 683 VWRWHF 689

RESULT 46
 C82715
 conserved hypothetical protein XF1178 [imported] - Xylella fastidiosa (strain 9a5c)
 C;Species: Xylella fastidiosa
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C;Accession: C82715
 R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A;Reference number: A82515; MUID:20365717; PMID:10910347
 A;Note: for a complete list of authors see reference number A59328 below
 A;Accession: C82715
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-106 <SIM>
 A;Cross-references: GB:AE003951; GB:AE003849; NID:99106131; PIDN:AAF83988.1; GSPDB:GN00:
 A;Experimental source: strain 9a5c
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; i
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, i
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohr
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.B.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, I
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.J
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasal
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; i
 A;Reference number: A59328
 A;Contents: annotation
 C;Genetics:

A:Gene: XF1178

Query Match 73.3%; Score 33; DB 2; Length 106;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
| | | | |
Db 5 WSRWYF 10

RESULT 47

C95132
hypothetical protein SP1142 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: C95132
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; He
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95132
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-136 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75252.1; PID:gl4972619; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1142

Query Match

Best Local Similarity 73.3%; Score 33; DB 2; Length 136;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 6
| | | | |
Db 19 WLKWF 24

RESULT 48

D84181
hypothetical protein Vng0208h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84181
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: D84181
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-149 <STO>
A:Cross-references: GB:AE004437; NID:g10579852; PIDN:AAG18816.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG0208H

Query Match

Best Local Similarity 73.3%; Score 33; DB 2; Length 149;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
| | | | |
Db 111 WLKWH 115

RESULT 49

T27980

hypothetical protein ZK678.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T27980
R;Kershaw, J.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z20449
A:Accession: T27980
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-221 <WIL>
A:Cross-references: EMBL:Z79605; PIDN:CAB01907.1; GSPDB:GN00028; CESP:ZK678.6
A:Experimental source: clone ZK678

C:Genetics:

A:Gene: CESP:ZK678.6

A:Map position: X

A:Introns: 63/3; 135/3

Query Match

Best Local Similarity 73.3%; Score 33; DB 2; Length 221;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
| | | | |
Db 137 WSRWYF 142

RESULT 50

A71020
hypothetical protein PH1454 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: A71020
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71020
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-241 <KAW>
A:Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30561.1; PID:g3257878
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1454

Query Match

Best Local Similarity 73.3%; Score 33; DB 2; Length 241;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWHF 6
| | | | |
Db 193 WQWTF 198

Search completed: June 10, 2004, 10:51:16

Job time : 20.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:36:54 ; Search time 8 Seconds
(without alignments)
39.053 Million cell updates/sec

Title: US-09-912-414-2

Perfect score: 45

Sequence: 1 WVRWHF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	42	93.3	479	1	CATA_PSEPU
2	40	88.9	443	1	FD6C_BRANA
3	40	88.9	448	1	FD6C_ARATH
4	39	86.7	475	1	ZCH5_HUMAN
5	39	86.7	482	1	CATA_ONCVE
6	38	84.4	822	1	TRKE_HUMAN
7	36	80.0	320	1	XERC_SYNP7
8	36	80.0	424	1	FD6C_SOYEN
9	36	80.0	788	1	BCSB_XANAC
10	36	80.0	1169	1	SUV3_DROME
11	35	77.8	226	1	YC07_METJA
12	35	77.8	254	1	HIS4_SYNP7
13	35	77.8	718	1	PLSB_CAEEL
14	34	75.6	148	1	RS12_METJA
15	34	75.6	316	1	CB1_ECOLI
16	34	75.6	316	1	CB1_KLEAE
17	34	75.6	358	1	MTCL_CAUCR
18	34	75.6	372	1	GP45_HUMAN
19	34	75.6	373	1	GP45_MOUSE
20	34	75.6	495	1	AOFN_ASPNG
21	34	75.6	799	1	AFSK_STRCO
22	34	75.6	807	1	AFSK_STRGR
23	34	75.6	1698	1	CUL7_HUMAN
24	33.5	74.4	1511	1	PDRC_YEAST
25	33	73.3	240	1	PSD_ZYMMO
26	33	73.3	244	1	YU68_CAEEL
27	33	73.3	376	1	MTS1_RHIME
28	33	73.3	377	1	MTB1_BRUAB
29	33	73.3	396	1	BCR_ECOLI
30	33	73.3	439	1	Y412_ARATH
31	33	73.3	478	1	CATA_LACSK
32	33	73.3	483	1	BCA_STRVL
33	33	73.3	502	1	CATA_TOXGO

34	33	73.3	503	1	CATA_MICLU	P29422 micrococcus
35	33	73.3	504	1	SIK1_YEAST	Q12460 saccharomyc
36	33	73.3	511	1	NOP5_YEAST	Q12499 saccharomyc
37	33	73.3	529	1	NOP5_HUMAN	Q9V2x3 homo sapien
38	33	73.3	534	1	NOP5_RAT	Q9QZ86 rattus norv
39	33	73.3	614	1	YG3F_YEAST	P53283 saccharomyc
40	33	73.3	821	1	TRKB_MOUSE	P15209 mus musculus
41	33	73.3	821	1	TRKB_RAT	Q63604 rattus norv
42	33	73.3	1015	1	FDNG_ECOLI	P24183 escherichia
43	33	73.3	1016	1	FDNG_ECOLI	P32176 escherichia
44	32	71.1	126	1	CRCB_FALSO	Q8XZr2 ralstonia s
45	32	71.1	129	1	YHW2_YEAST	P38857 saccharomyc
46	32	71.1	146	1	YG56_YEAST	P53311 saccharomyc
47	32	71.1	228	1	LIPB_THETN	Q8R9e0 thermoanaer
48	32	71.1	260	1	COBS_SYNY3	O55714 synechocyst
49	32	71.1	287	1	YAHE_ECOLI	P77297 escherichia
50	32	71.1	324	1	YJLA_BACSU	O34428 bacillus su
51	32	71.1	386	1	RL4_URECA	F49165 urechis cau
52	32	71.1	395	1	DLTB_BACSU	P39580 bacillus su
53	32	71.1	396	1	RL4A_XENLA	P08429 xenopus lae
54	32	71.1	396	1	RL4B_XENLA	P02385 xenopus lae
55	32	71.1	406	1	RL4B_ARATH	Q98f40 arabidopsis
56	32	71.1	407	1	RL4A_ARATH	P49691 arabidopsis
57	32	71.1	408	1	RL4_PRUAR	Q9Xf97 prunus arme
58	32	71.1	419	1	RL4_MOUSE	Q9D8e6 mus musculus
59	32	71.1	420	1	RL4_CANFA	Q28346 canis famil
60	32	71.1	421	1	RL4_RAT	P50878 rattus norv
61	32	71.1	422	1	TERL_BPSPP	P54308 bacteriopho
62	32	71.1	427	1	RL4_HUMAN	P36578 homo sapien
63	32	71.1	447	1	VIPIR_CARAU	Q90308 carassius a
64	32	71.1	457	1	VIPIR_HUMAN	P32241 homo sapien
65	32	71.1	457	1	VIPIR_PELG	Q91085 meleagris g
66	32	71.1	458	1	VIPIR_PIG	Q28992 sus scrofa
67	32	71.1	459	1	VIPIR_MOUSE	P97751 mus musculus
68	32	71.1	459	1	VIPIR_MOUSE	P97751 mus musculus
69	32	71.1	482	1	CATA_BACSU	P30083 rattus norv
70	32	71.1	482	1	CATA_PSEAE	P26901 bacillus su
71	32	71.1	482	1	CATA_VIBFI	O52762 pseudomonas
72	32	71.1	484	1	CATA_PROMI	O68146 vibrio fisc
73	32	71.1	485	1	ALGI_PSEPK	P42321 proteus mir
74	32	71.1	494	1	CATA_RHIME	Q88nd2 pseudomonas
75	32	71.1	495	1	ALGI_PSEFL	P95631 rhizobium m
76	32	71.1	498	1	CATA_BRUME	P59789 pseudomonas
77	32	71.1	498	1	CATA_BRUME	Q59170 bruceella me
78	32	71.1	498	1	CATA_BRUSE	O8fwu0 bruceella su
79	32	71.1	499	1	ALGI_AZOVI	O52196 azotobacter
80	32	71.1	500	1	CATA_NEIGO	Q59602 neisseria g
81	32	71.1	505	1	CATA_HELPJ	Q92kx5 helicobacte
82	32	71.1	505	1	CATA_HELPJ	P77872 helicobacte
83	32	71.1	505	1	CATA_METEA	Q93662 methanosarc
84	32	71.1	507	1	CATA_CAMTE	Q59296 campylobact
85	32	71.1	508	1	C983_ARATH	O22203 arabidopsis
86	32	71.1	508	1	CATA_HAEIN	P44390 haemophilus
87	32	71.1	509	1	C982_SOYBN	O48922 glycine max
88	32	71.1	518	1	ALGI_PSESM	Q887q6 pseudomonas
89	32	71.1	520	1	ALGI_PSEAE	Q51392 pseudomonas
90	32	71.1	1028	1	FXDG_HAEIN	P46448 haemophilus
91	32	71.1	1034	1	BGAL_KLEPN	P06219 klebsiella
92	32	71.1	1193	1	DPOL_ADE04	P17452 pasteurella
93	31.5	70.0	956	1	XDHD_PASMU	P87503 human adeno
94	31.5	70.0	956	1	XDHD_ECO57	Q8xd64 escherichia
95	31	68.9	119	1	V5IS_REOVL	Q46814 escherichia
96	31	68.9	170	1	CBP4_YEAST	P07938 reovirus (t
97	31	68.9	195	1	TNF7_MOUSE	Q55237 mus musculus
98	31	68.9	223	1	PGC2_HUMAN	P37267 saccharomyc
99	31	68.9	295	1	MORA_PSEPU	O52327 mus musculus
100	31	68.9	307	1	TR41_HUMAN	O15173 homo sapien
101	31	68.9	307	1	TR59_HUMAN	Q02198 pseudomonas
102	31	68.9	308	1	TR59_MOUSE	P59536 homo sapien
103	31	68.9	308	1	T2RC_MOUSE	P59550 homo sapien
104	31	68.9	308	1	T2RC_RAT	P59532 mus musculus
105	31	68.9	310	1	ARAC_ERWCH	Q91ke7 rattus norv
106	31	68.9	409	1	POBA_PSEPS	P07642 erwinia chr
106	31	68.9	418	1	RHAA_BACHD	Q52185 pseudomonas
						Q9kcl9 bacillus ha

107	31	68.9	436	1	ACHX ONCVO	P54247 onchocerca	180	30	66.7	550	1	PM21 LYCES	P09507 lycopersico
108	31	68.9	452	1	AAMP HUMAN	Q13685 homo sapien	181	30	66.7	611	1	SNF1_CANGA	Q00372 candida gla
109	31	68.9	475	1	ETS6 DROME	P29776 drosophila	182	30	66.7	619	1	SNF1_CANTR	Q94168 candida tro
110	31	68.9	560	1	NMB_HUMAN	Q14956 homo sapien	183	30	66.7	620	1	SNF1_CANAL	P52497 candida alb
111	31	68.9	614	1	PEOB_STIN3	P73182 synchocyst	184	30	66.7	633	1	SNF1_YEAST	P06782 saccharomyc
112	31	68.9	640	1	Y4CD RHIN3	P55386 rhizobium s	185	30	66.7	634	1	SNF1_RHIME	P15715 rhizobium m
113	31	68.9	754	1	YCAI_ECOLI	P37443 escherichia	186	30	66.7	663	1	PD13 SHEEP	O02849 ovis aries
114	31	68.9	856	1	ENV_FIVPE	P16090 feline immu	187	30	66.7	702	1	FOXA_SALTU	Q56145 salmonella
115	31	68.9	863	1	RPOC_NEPOL	Q9T105 neproselmi	188	30	66.7	805	1	YC26_METJA	Q58623 methanococc
116	31	68.9	901	1	ODPI MYCTU	Q10504 mycobacteri	189	30	66.7	815	1	YBQO_ECOLI	P75750 escherichia
117	31	68.9	1266	1	NGCA_CHICK	Q03696 gallus gall	190	30	66.7	1046	1	POL_FENVI	P31792 feline endo
118	31	68.9	3341	1	POLG_MCF4	P33515 m genome po	191	30	66.7	1257	1	ERE2 RAT	P06494 rattus norv
119	30.5	67.8	641	1	DHSA SCHPO	Q9UTJ7 schizosacch	192	30	66.7	1281	1	YLB5_CAEEL	Q9K472 bacillus ha
120	30.5	67.8	646	1	DHSA_CAEEL	Q09508 caenorhabdi	193	30	66.7	1433	1	DPO3_BACHD	O42690 candida alb
121	30.5	67.8	664	1	DHSA_HUMAN	P31040 homo sapien	194	30	66.7	1501	1	CDR3_CANAL	Q77086 drosophila
122	30.5	67.8	665	1	DHSA_BOVIN	P31039 bos taurus	195	30	66.7	1571	1	C3G_DROME	Q28107 bos taurus
123	30	66.7	67	1	YPB1_NPVLD	P36866 lymantria d	196	30	66.7	2211	1	FAS_BOVIN	P12259 homo sapien
124	30	66.7	140	1	LYSA_DROME	P37157 drosophila	197	30	66.7	2224	1	FAS_HUMAN	Q9GLP1 sus scrofa
125	30	66.7	140	1	LYSB_DROME	Q08694 drosophila	198	30	66.7	2258	1	FAS_PIG	P07720 t genome po
126	30	66.7	140	1	LYSE_DROME	P37159 drosophila	199	30	66.7	3412	1	POLG_TBEVS	P29837 t genome po
127	30	66.7	140	1	LYSS_DROME	P37160 drosophila	200	30	66.7	3414	1	POLG_LANVT	Q01299 t genome po
128	30	66.7	142	1	LYSX_DROME	P37161 drosophila	201	30	66.7	3414	1	POLG_TBEVH	P14336 t genome po
129	30	66.7	164	1	LSPA_ECO57	Q8XA48 escherichia	202	30	66.7	3414	1	POLG_TBEVW	Q04538 t genome po
130	30	66.7	164	1	LSPA_ECOL6	Q8FLB6 escherichia	203	30	66.7	3415	1	POLG_POWVL	P98157 gallus gall
131	30	66.7	164	1	LSPA_ECOL1	P00804 escherichia	204	30	66.7	4543	1	LRP1_CHICK	Q07954 homo sapien
132	30	66.7	165	1	LSPA_ENTAE	P13514 enterobacte	205	30	66.7	4544	1	LRP1_HUMAN	P21878 bacillus st
133	30	66.7	166	1	LSPA_SALTU	Q829N1 salmonella	206	29	64.4	95	1	YPGA_BACST	Q05178 homo sapien
134	30	66.7	166	1	LSPA_SALTU	Q829N1 salmonella	207	29	64.4	105	1	NIGM_HUMAN	Q02374 bos taurus
135	30	66.7	169	1	LSPA_VERPE	Q8XIL9 yersinia pe	208	29	64.4	108	1	NYGM_BOVIN	Q32211 bacillus su
136	30	66.7	172	1	Y819_PASMU	Q9CMK4 pasteurella	209	29	64.4	161	1	YVGO_BACSU	P95211 mycobacteri
137	30	66.7	215	1	UR2R_BOVIN	P49220 bos taurus	210	29	64.4	181	1	AAC2_MYCTU	Q92KE6 helicobacte
138	30	66.7	216	1	VP2_CAV26	P54092 chicken ane	211	29	64.4	213	1	SODF_HELPJ	Q97CW0 clostridium
139	30	66.7	216	1	VP2_CAV82	Q91511 chicken ane	212	29	64.4	254	1	OXAA_CLOAB	P75851 escherichia
140	30	66.7	216	1	VP2_CAVC1	P75407 mycoplasma	213	29	64.4	263	1	SSUC_ECOLI	Q9N0P9 bos taurus
141	30	66.7	229	1	YD74_MYCPN	Q29844 archaeoglob	214	29	64.4	313	1	PIMI_BOVIN	Q95110 felis silve
142	30	66.7	231	1	PCRB_ARCTU	Q8LUA0 homo sapien	215	29	64.4	313	1	PIMI_HUMAN	P11309 homo sapien
143	30	66.7	241	1	WED8_HUMAN	Q8XAQ3 escherichia	216	29	64.4	313	1	PIMI_MOUSE	P06803 mus musculu
144	30	66.7	246	1	WECG_ECO57	Q8FBP7 escherichia	217	29	64.4	314	1	MTM1_HAEP	P26794 rattus norv
145	30	66.7	246	1	WECG_ECOL6	P27836 escherichia	218	29	64.4	314	1	ML34_ARATH	P29538 haemophilus
146	30	66.7	246	1	WECG_ECOL1	Q82397 salmonella	219	29	64.4	316	1	LM34_COTJA	Q9SK77 arabidopsis
147	30	66.7	246	1	WECG_SALTU	P37457 salmonella	220	29	64.4	316	1	PIM3_COTJA	Q9PUB5 coturnix co
148	30	66.7	246	1	UR1E_LEPIN	Q8XJ33 leptospira	221	29	64.4	316	1	PIM3_MOUSE	Q86V86 homo sapien
149	30	66.7	249	1	YBDC_YEAST	P38197 saccharomyc	222	29	64.4	316	1	PIM3_MOUSE	P58750 mus musculu
150	30	66.7	257	1	YBDC_YEAST	Q04470 bacillus th	223	29	64.4	343	1	Y098_GVCL	Q70444 rattus norv
151	30	66.7	259	1	CYAA_BACTY	Q94903 homo sapien	224	29	64.4	343	1	ARGC_PSEAE	P41729 cryptophleb
152	30	66.7	274	1	POSC_MOUSE	Q2ZEL6 pseudomonas	225	29	64.4	344	1	YA03_TREPA	Q91549 pseudomonas
153	30	66.7	275	1	LEP4_PSEST	P22610 pseudomonas	226	29	64.4	345	1	YA03_TREPA	O83968 treponema p
154	30	66.7	289	1	LEP4_PSEST	Q44185 agrobacteri	227	29	64.4	345	1	RL4B_SCHPO	P35679 schizosacch
155	30	66.7	290	1	LEP4_PSEAE	Q91822 xenopus lae	228	29	64.4	363	1	RL4B_SCHPO	Q9P784 schizosacch
156	30	66.7	304	1	DCAS_AGRUT	Q11058 mycobacteri	229	29	64.4	363	1	O33B_DROME	P81915 drosophila
157	30	66.7	323	1	PIM3_XENLA	P08740 chlamydomon	230	29	64.4	379	1	T10B_MOUSE	Q9GZM4 mus musculu
158	30	66.7	372	1	YC60_MYCTU	Q50695 mycobacteri	231	29	64.4	432	1	FIBG_PETMA	P04115 petromyzon
159	30	66.7	382	1	NU2M_CHLRE	P41929 yarrowia li	232	29	64.4	433	1	YBBY_ECOLI	P77328 escherichia
160	30	66.7	388	1	LMV61_MYCTU	P45064 haemophilus	233	29	64.4	440	1	SCRC_HUMAN	P47872 homo sapien
161	30	66.7	391	1	LYC1_YARLI	P55480 rhizobium s	234	29	64.4	445	1	YF71_PASMU	O46502 cryptotlagus
162	30	66.7	394	1	FTSW_HABIN	Q912A6 mus musculu	235	29	64.4	446	1	CATA_DICDI	Q9CKN9 pasteurallu
163	30	66.7	403	1	Y4HM_RHISN	O8WTH5 homo sapien	236	29	64.4	496	1	YE91_HELPY	O77229 dictyosteli
164	30	66.7	426	1	YA35_SCHPO	P31474 escherichia	237	29	64.4	533	1	FIXN_AGR77	O26024 helicobacte
165	30	66.7	436	1	SOC7_MOUSE	P46216 trichomonas	238	29	64.4	539	1	FIXN_RHIME	P98055 agrobacteri
166	30	66.7	440	1	SOC7_MOUSE	Q9Y264 homo sapien	239	29	64.4	539	1	FIXN_RHIME	Q05072 rhizobium m
167	30	66.7	475	1	YVIO_ECOLI	Q57541 pseudomonas	240	29	64.4	549	1	CAH_DUNSA	Q03073 bradyrhizob
168	30	66.7	489	1	SVY_TRIVA	Q9Y264 homo sapien	241	29	64.4	589	1	YGL3_YEAST	P54212 dunaliella
169	30	66.7	503	1	AGP4_HUMAN	Q46087 chelatobact	242	29	64.4	668	1	RPOC_WHEAT	P53135 saccharomyc
170	30	66.7	507	1	TR6A_PSEAE	Q9WVH6 mus musculu	243	29	64.4	686	1	CHEA_RHOSH	Q9XPS8 triticum ae
171	30	66.7	508	1	MLO3_ARATH	P38840 saccharomyc	244	29	64.4	762	1	ACV_BOVIN	O46406 bos taurus
172	30	66.7	508	1	TRAT_CHEHE	O06127 mycobacteri	245	29	64.4	763	1	AOC3_HUMAN	Q16853 homo sapien
173	30	66.7	509	1	AGP4_MOUSE	P18396 rhizobium m	246	29	64.4	765	1	AOC3_MOUSE	O70423 mus musculu
174	30	66.7	511	1	FAST_MOUSE	P57210 buchnera ap	247	29	64.4	771	1	TMEB_HUMAN	Q9HCN3 homo sapien
175	30	66.7	513	1	AROS_YEAST	Q14296 homo sapien	248	29	64.4	803	1	OPGH_PSESY	P23622 neurospora
176	30	66.7	516	1	TRPE_MOUSE		249	29	64.4	826	1	RIR1_EBV	P03190 epstein-bar
177	30	66.7	524	1	FIXG_RHIME		250	29	64.4				
178	30	66.7	547	1	SYM_BUCAI		251	29	64.4				
179	30	66.7	549	1	FAST_HUMAN		252	29	64.4				

253	29	64.4	850	1	STB2_YEAST	P46679	saccharomyc	326	28	62.2	347	1	ND2M_DUGDU	Q8w9n5	dugong dugo
254	29	64.4	857	1	OPGH_PSEPK	Q88d04	pseudomonas	327	28	62.2	349	1	TRM1_ARCFU	Q29443	archaeoglob
255	29	64.4	859	1	OPGH_PSEPK	Q87uy1	pseudomonas	328	28	62.2	355	1	TSY3_HUMAN	Q9h489	homo sapien
256	29	64.4	861	1	OPGH_PSEAE	Q9h486	pseudomonas	329	28	62.2	363	1	YGEW_ECOLI	Q46803	escherichia
257	29	64.4	918	1	PR3_YEAST	P27801	saccharomyc	330	28	62.2	366	1	YACL_BACSU	Q06754	bacillus su
258	29	64.4	974	1	CL15_YEAST	P27636	saccharomyc	331	28	62.2	370	1	NAM2_HUMAN	Q15035	homo sapien
259	29	64.4	982	1	POL_HTLV2	P03363	human t-cel	332	28	62.2	380	1	NER2_HUMAN	Q9y3r4	homo sapien
260	29	64.4	1002	1	EPB5_CHICK	Q07497	gallus gall	333	28	62.2	380	1	YCK7_ODOSI	P49833	odontella s
261	29	64.4	1018	1	VGNM_BPMV	P23009	bean-pod mo	334	28	62.2	382	1	YCAD_ECOLI	P21503	escherichia
262	29	64.4	1198	1	UB42_HUMAN	Q8h9j4	homo sapien	335	28	62.2	385	1	ST22_YEAST	P25604	saccharomyc
263	29	64.4	1520	1	ACFD_VIBCH	Q9tkq4	vibrio chol	336	28	62.2	388	1	NH16_CAEEL	Q27521	caenorhabdi
264	29	64.4	3487	1	CSM2_HUMAN	Q7z408	homo sapien	337	28	62.2	396	1	YGEW_ECOLI	Q8x6c0	escherichia
265	29	64.4	3511	1	MY15_MOUSE	Q9qz24	mus musculus	338	28	62.2	396	1	YGEW_ECOLI	Q8feg1	escherichia
266	29	64.4	3564	1	CSM1_MOUSE	Q96p27	homo sapien	339	28	62.2	398	1	OPRM_MOUSE	P42866	mus musculus
267	29	64.4	3670	1	CSM1_HUMAN	Q96p27	homo sapien	340	28	62.2	398	1	OPRM_MOUSE	P33532	rattus norv
268	28.5	63.3	512	1	NIFK_HUMAN	Q7z407	homo sapien	341	28	62.2	400	1	OPRM_HUMAN	P35373	homo sapien
269	28.5	63.3	933	1	GIND_AZORR	P00468	anabaena sp	342	28	62.2	400	1	OPRM_HUMAN	Q9myw9	macaca mula
270	28	62.2	70	1	Y3C4_STRCO	Q8rqd1	azospirillum	343	28	62.2	401	1	OPRM_BOVIN	P79350	bos taurus
271	28	62.2	83	1	Y13L_BPT4	P33866	streptomyce	344	28	62.2	401	1	OPRM_PIG	Q95247	sus scrofa
272	28	62.2	83	1	Y598_MYCLE	P39505	bacteriophage	345	28	62.2	407	1	NH86_CAEEL	Q965w2	caenorhabdi
273	28	62.2	96	1	VFX_HV2D1	Q33024	mycobacteri	346	28	62.2	424	1	RHAA_BACSU	O05264	bacillus su
274	28	62.2	112	1	TVCL_MOUSE	P17760	human immun	347	28	62.2	425	1	YH29_PVRAB	O8uxx9	pyrococcus
275	28	62.2	135	1	TVCL_MOUSE	P17760	human immun	348	28	62.2	425	1	YH29_PVRAB	O8uxx9	pyrococcus
276	28	62.2	149	1	GLB1_MORMR	P01740	mus musculus	349	28	62.2	427	1	YF09_MYCPN	O59545	pyrococcus
277	28	62.2	149	1	GLB1_PETMA	P21197	mordacia mo	350	28	62.2	434	1	SY8_RHILLO	P75277	mycoplasma
278	28	62.2	149	1	GLB2_MORMR	P21198	mordacia mo	351	28	62.2	440	1	AM3A_ORYSA	Q981c8	rhizobium 1
279	28	62.2	149	1	GLB2_PETMA	Q91913	petromyzon	352	28	62.2	448	1	SYC_LACLA	P27932	oryza sativ
280	28	62.2	149	1	GLB3_MORMR	P09968	petromyzon	353	28	62.2	449	1	SCRC_RAT	Q9ce10	lactococcus
281	28	62.2	149	1	GLB3_PETMA	P09968	petromyzon	354	28	62.2	450	1	YOE9_PSEAE	P23811	rattus norv
282	28	62.2	149	1	GLB5_PETMA	P02208	petromyzon	355	28	62.2	450	1	YOE9_PSEAE	Q9hxy3	pseudomonas
283	28	62.2	149	1	GLB5_PETMA	P02207	lampetra fl	356	28	62.2	453	1	NH12_CAEEL	Q8y2h7	anabaena sp
284	28	62.2	168	1	LSPA_VIBVU	Q8des8	vibrio vuln	357	28	62.2	453	1	NH12_CAEEL	Q21701	caenorhabdi
285	28	62.2	169	1	LSPA_PSEAE	Q9hvm5	pseudomonas	358	28	62.2	458	1	YF10_MYCPN	P75276	mycoplasma
286	28	62.2	169	1	LSPA_PSEAE	Q9hvm5	pseudomonas	359	28	62.2	462	1	SYTM_YEAST	P07236	saccharomyc
287	28	62.2	169	1	LSPA_VIBPA	P36058	saccharomyc	360	28	62.2	464	1	MNTH_LACPL	Q8gh68	lactobacill
288	28	62.2	170	1	LSPA_PSEFL	P17942	pseudomonas	361	28	62.2	465	1	NPT1_MOUSE	Q61983	mus musculus
289	28	62.2	171	1	LSPA_PSEPK	Q88q91	pseudomonas	362	28	62.2	465	1	NPT1_RABIT	Q28722	o renal sod
290	28	62.2	171	1	LSPA_PSEPK	Q9ku46	vibrio chol	363	28	62.2	466	1	GUN5_THEFU	Q62795	rattus norv
291	28	62.2	172	1	LSPA_PSEPK	Q9ku46	vibrio chol	364	28	62.2	469	1	RDXA_RHOSH	Q01786	thermomonos
292	28	62.2	173	1	YEOJ_ECOLI	P31131	escherichia	365	28	62.2	472	1	PPB_ESCFE	Q01854	rhodobacter
293	28	62.2	178	1	LSPA_PSEPM	Q88963	pseudomonas	366	28	62.2	477	1	RDXB_RHOSH	P21948	escherichia
294	28	62.2	189	1	DSBB_VIBPA	Q87n03	vibrio para	367	28	62.2	483	1	SHT3_RAT	P54932	rhodobacter
295	28	62.2	197	1	VP21_AMCV	P15961	artichoke m	368	28	62.2	486	1	YDR1_ECOLI	P35563	rattus norv
296	28	62.2	197	1	DSBE_XLIPA	Q9pan4	xylella fas	369	28	62.2	487	1	SHT3_MOUSE	P77260	escherichia
297	28	62.2	197	1	DSBE_XLIFT	Q87bh3	xylella fas	370	28	62.2	489	1	Y909_VIBPA	P23379	mus musculus
298	28	62.2	208	1	YHNN_ECOLI	P37616	escherichia	371	28	62.2	489	1	YJ31_VIBCH	Q9kdr7	vibrio chol
299	28	62.2	208	1	YQED_BACSU	P54449	bacillus su	372	28	62.2	490	1	SHT3_CAVPO	O70212	cavia porce
300	28	62.2	213	1	SODF_HILPY	P43312	helicobacte	373	28	62.2	493	1	C6AD_DROME	Q9v4u9	drosofila
301	28	62.2	218	1	Y556_SNY3	P52056	synchocyst	374	28	62.2	504	1	YC03_KLEPN	Q48449	klebsiella
302	28	62.2	233	1	YPI1_VIBAL	P52055	vibrio algi	375	28	62.2	509	1	C6A1_MUSDO	P13527	musca domes
303	28	62.2	234	1	S120_YEAST	P09931	saccharomyc	376	28	62.2	512	1	CATA_SCHPO	P55306	schizosacch
304	28	62.2	251	1	UBIE_ECOLI	Q10423	schizosacch	377	28	62.2	514	1	CBH2_SCHPO	O60108	schizosacch
305	28	62.2	258	1	CCEI_SCHPO	P59531	homo sapien	378	28	62.2	517	1	COX1_SULAC	O14423	schizosacch
306	28	62.2	264	1	T2RC_HUMAN	P59531	homo sapien	379	28	62.2	522	1	ABP1_SCHPO	P43777	schizosacch
307	28	62.2	286	1	PURU_CORSP	Q9nyv7	homo sapien	380	28	62.2	532	1	COX1_RHOC	P98004	sulfolobus
308	28	62.2	291	1	T2RG_HUMAN	Q8txs4	methanopyru	381	28	62.2	533	1	YB91_HELPF	P43777	schizosacch
309	28	62.2	296	1	PRIS_METKA	P25325	homo sapien	382	28	62.2	533	1	YB91_HELPF	P98059	rhodobacte
310	28	62.2	296	1	THIM_HUMAN	Q9nyv7	homo sapien	383	28	62.2	546	1	PTH2_RAT	Q9xj08	helicobacte
311	28	62.2	296	1	THIM_RAT	P32143	rattus norv	384	28	62.2	551	1	FIXN_AZOCA	P70555	rattus norv
312	28	62.2	298	1	YIHW_ECOLI	P32143	rattus norv	385	28	62.2	559	1	PP02_MOUSE	P98055	azorhizobiu
313	28	62.2	308	1	BEL1_SGVI	P29169	simian foam	386	28	62.2	571	1	CAN1_CANAL	O88554	mus musculus
314	28	62.2	309	1	PIP_MYCPN	P75092	mycoplasma	387	28	62.2	583	1	PP02_MOUSE	P43059	candida alb
315	28	62.2	310	1	DNUL_MYCGE	Q47248	mycoplasma	388	28	62.2	583	1	SIS_HUMAN	Q9ugn5	homo sapien
316	28	62.2	312	1	FGU1_HUMAN	Q08830	homo sapien	389	28	62.2	608	1	PLRL_MOUSE	P08842	homo sapien
317	28	62.2	312	1	YCBK_BACSU	P42423	bacillus su	390	28	62.2	610	1	PLRL_MOUSE	Q08501	mus musculus
318	28	62.2	320	1	ASPE_SPOFR	O02467	spodoptera	391	28	62.2	616	1	PLRL_RABIT	P05710	rattus norv
319	28	62.2	322	1	ASPE_SPOFR	P76215	escherichia	392	28	62.2	624	1	STS_MOUSE	P14787	oryctolagus
320	28	62.2	325	1	UL76_HQWVA	P16725	human cytom	393	28	62.2	637	1	SKB1_HUMAN	P50427	mus musculus
321	28	62.2	335	1	PISD_CAEEL	Q10949	caenorhabdi	394	28	62.2	637	1	SKB1_MOUSE	O14744	homo sapien
322	28	62.2	335	1	CD1D_HUMAN	P15813	homo sapien	395	28	62.2	666	1	Y032_MYCGE	Q8cig8	mus musculus
323	28	62.2	341	1	PDXA_RALSO	P58714	raletonia s	396	28	62.2	666	1	YA35_MYCPN	P47578	mycoplasma
324	28	62.2	338	1	ETV7_HUMAN	Q9y603	homo sapien	397	28	62.2	671	1	AM01_ASPNG	P75079	mycoplasma
325	28	62.2	344	1	NLBP_ECOLI	P21167	escherichia	398	28	62.2	673	1	YA36_MYCPN	Q12556	aspergillus
														P75078	mycoplasma

399	28	62.2	674	1	CWF4_SCHPO	P87312	schizosacch	153	1	LSPA_WIGBR	Q8d2r1	wigleswort
400	28	62.2	677	1	RPOC_SPTOL	P11705	spinacia ol	154	1	ELYS_HALSO	Q01383	haliotis so
401	28	62.2	679	1	RPOC_OENHO	Q5mtm4	oenothera h	162	1	YBUO_ECOLI	P75816	escherichia
402	28	62.2	680	1	RPOC_ARATH	P57673	arabidopsis	172	1	YKGE_BACSU	P49855	bacillus su
403	28	62.2	681	1	RPOC_TORAC	P12116	nicotiana t	179	1	RLS_BUCAI	P57579	buchnera ap
404	28	62.2	682	1	RPOC_LOTJA	Q9bbs8	lotus japon	187	1	NIP3_MOUSE	O55003	mus musculus
405	28	62.2	682	1	RPOC_ORYZA	P12092	oryza sativ	191	1	PORC_THEMEA	O05650	thermotoga
406	28	62.2	683	1	RPOC_MAIZE	P16024	zea mays (m	194	1	NIP3_HUMAN	Q12983	homo sapien
407	28	62.2	684	1	RPOC_MARPO	P06273	marichantia	200	1	SP24_BOVIN	Q27967	bos taurus
408	28	62.2	688	1	RPOC_SINAL	P46819	sinapis alb	202	1	Y549_BUCAI	Q26783	buchnera ap
409	28	62.2	691	1	PK9_RAT	P59996	rattus norv	202	1	NURM_TRYBB	P37027	escherichia
410	28	62.2	694	1	PK9_MOUSE	Q80w65	mus musculus	209	1	IDI_AGRRH	Q9kwd1	agrobacteri
411	28	62.2	696	1	PK9_PINTH	P52733	pinus thumb	210	1	SSH5_YEAST	Q03446	saccharomyc
412	28	62.2	736	1	GLGB_PSEPK	Q88fn1	pseudomonas	221	1	FLIH_BUCAP	Q8ka43	buchnera ap
413	28	62.2	747	1	POT1_HUMAN	Q9y6a1	homo sapien	223	1	VMAT_SVCV	P04888	spring vire
414	28	62.2	747	1	YME1_YEAST	P32795	saccharomyc	227	1	US08_HCMVA	P09730	human cytom
415	28	62.2	769	1	TME8_MOUSE	Q61139	mus musculus	231	1	GPMA_BUCAI	P57390	buchnera ap
416	28	62.2	770	1	PK7_MOUSE	Q61139	mus musculus	233	1	Y53L_SYNT3	P72583	synchocyst
417	28	62.2	783	1	PK7_RAT	Q62849	rattus norv	242	1	YN82_VIBCH	Q87qk1	vibrio chol
418	28	62.2	785	1	PK7_HUMAN	Q16549	homo sapien	248	1	LFXH_RALSO	Q8xyy9	raistonia s
419	28	62.2	836	1	RPOC_CHLAV	P56300	chlorella v	250	1	YK21_YEAST	O94611	schizosacch
420	28	62.2	838	1	KFC2_HUMAN	Q96ac6	homo sapien	251	1	LP4_BOMMO	P09337	bombyx mori
421	28	62.2	859	1	ST7_HUMAN	Q9y561	homo sapien	254	1	RPSD_BACSU	P10726	bacillus su
422	28	62.2	862	1	OPGH_RALSO	O13944	schizosacch	254	1	YABI_ECOLI	P30149	escherichia
423	28	62.2	945	1	YEH1_SCHPO	P33231	lycopersico	260	1	EBS1_YEAST	P17261	saccharomyc
424	28	62.2	960	1	PT09_YEAST	P32522	saccharomyc	260	1	RFBA_MYXXA	Q50862	myxococcus
425	28	62.2	965	1	VP41_YEAST	P93043	arabidopsis	27	60.0			
426	28	62.2	976	1	VP41_ARATH	F25218	equine heip	27	60.0			
427	28	62.2	979	1	VGLB_HSVEL	P33537	neurospora	27	60.0			
428	28	62.2	1021	1	DPOM_NEUCR	P38850	saccharomyc	27	60.0			
429	28	62.2	1070	1	YHV4_YEAST	Q9heh1	neurospora	27	60.0			
430	28	62.2	1093	1	RNT1_NEUCR	P33459	caprine art	27	60.0			
431	28	62.2	1109	1	POL_CAEVC	P49455	haemophilus	27	60.0			
432	28	62.2	1121	1	EX5C_HAEIN	Q43052	schizosacch	27	60.0			
433	28	62.2	1150	1	RGAI_SCHPO	Q9rkby	streptomyce	27	60.0			
434	28	62.2	1171	1	TR12_STRCO	P33346	escherichia	27	60.0			
435	28	62.2	1210	1	YEH1_ECOLI	Q12769	homo sapien	27	60.0			
436	28	62.2	1316	1	N160_HUMAN	Q920w3	mus musculus	27	60.0			
437	28	62.2	1402	1	N160_MOUSE	Q04351	clostridium	27	60.0			
438	28	62.2	1498	1	Y1A9_CLOAB	Q46837	escherichia	27	60.0			
439	28	62.2	1520	1	ACFD_ECOLI	Q4766	homo sapien	27	60.0			
440	28	62.2	1595	1	LTBL_HUMAN	O53114	mycobacteri	27	60.0			
441	28	62.2	1609	1	CTPI_MYCLE	P47171	saccharomyc	27	60.0			
442	28	62.2	1609	1	YPSH_YEAST	Q00918	rattus norv	27	60.0			
443	28	62.2	1648	1	Y3H1_RAT	Q8cg19	mus musculus	27	60.0			
444	28	62.2	1712	1	LTBL_MOUSE	Q9ul36	homo sapien	27	60.0			
445	28	62.2	1713	1	LTBL_MOUSE	P39510	homo sapien	27	60.0			
446	28	62.2	1845	1	Z236_HUMAN	Q01901	p genome po	27	60.0			
447	28	62.2	1911	1	AT20_HUMAN	Q9nzj4	homo sapien	27	60.0			
448	28	62.2	3344	1	POLG_PSVH	Q9jlc8	mus musculus	27	60.0			
449	28	62.2	3829	1	SACS_HUMAN	P15969	agelenopsis	27	60.0			
450	28	62.2	3830	1	SACS_MOUSE	Q57860	methanococ	27	60.0			
451	27.5	61.1	112	1	TX1A_AGEAP	Q9m3j1	spinacia ol	27	60.0			
452	27.5	61.1	323	1	Y417_METJA	Q04442	synchococ	27	60.0			
453	27.5	61.1	323	1	CCSA_SPTOL	O52288	mycobacteri	27	60.0			
454	27.5	61.1	417	1	YAT1_SYNP6	P18328	mycobacteri	27	60.0			
455	27.5	61.1	1722	1	LY75_HUMAN	P08449	synchococ	27	60.0			
456	27	60.0	85	1	YGM2_DENAN	Q60449	homo sapien	27	60.0			
457	27	60.0	86	1	Y505_MYCPN	Q52288	mycobacteri	27	60.0			
458	27	60.0	101	1	HSP2_RAT	P18328	mycobacteri	27	60.0			
459	27	60.0	104	1	HSP2_RAT	P75147	mycoplasma	27	60.0			
460	27	60.0	106	1	YQCC_HAEIN	P11248	rattus norv	27	60.0			
461	27	60.0	108	1	XJJB_ECOLI	Q57152	haemophilus	27	60.0			
462	27	60.0	117	1	HV30_HUMAN	P18389	escherichia	27	60.0			
463	27	60.0	118	1	YAI9_SCHPO	P01776	homo sapien	27	60.0			
464	27	60.0	119	1	HV3M_HUMAN	Q09896	schizosacch	27	60.0			
465	27	60.0	119	1	HV3N_HUMAN	P01774	homo sapien	27	60.0			
466	27	60.0	119	1	HV3P_HUMAN	P01775	homo sapien	27	60.0			
467	27	60.0	123	1	Y477_MYCLE	P01777	homo sapien	27	60.0			
468	27	60.0	124	1	YH57_ARCFU	Q9cb44	mycobacteri	27	60.0			
469	27	60.0	145	1	PA21_LATIA	O28517	archaeoglob	27	60.0			
470	27	60.0	145	1	PA22_LATIA	P19000	laticauda l	27	60.0			
471	27	60.0	145	1	PA23_LATIA	Q8auu4	laticauda l	27	60.0			
	27	60.0	152	1	RL30_ARCFU	Q8auu3	laticauda l	27	60.0			
	27	60.0				O28375	archaeoglob	27	60.0			

ALIGNMENTS

RESULT 1

ID	CATA_PSEPU	STANDARD;	PRT;	479 AA.
AC	Q59714;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Catalase (EC 1.11.1.6).			
GN	KATA OR CATA.			
OS	Pseudomonas putida.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;			
OC	Pseudomonadaceae; Pseudomonas.			
[1]	NCBI_TaxID=303;			
[1]	SEQUENCE FROM N.A.			
RP	STRAIN=Corvallis;			
RC	MEDLINE=98019091; PubMed=9358059;			
EX	Kim Y.C., Miller C.D., Anderson A.J.;			
RA	"Identification of adjacent genes encoding the major catalase and a			
RT	bacterioferritin from the plant-beneficial bacterium Pseudomonas			
RL	putida.";			
CC	Gene 199:219-224(1997).			
CC	-!- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;			
CC	serves to protect cells from the toxic effects of hydrogen			
CC	peroxide.			
CC	-!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.			
CC	-!- COFACTOR: Heme group.			
CC	-!- ENZYME REGULATION: ACTIVATED BY PEROXIDE.			
CC	-!- SIMILARITY: Belongs to the catalase family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; U63511; AAB88219.1; -.			

DR HSP: P42321; 2CAE.
 DR InterPro: IPR002226; Catalase.
 DR Pfam: PF00199; catalase; 1.
 DR PRINTS: PR00067; CATALASE.
 DR PRODOM: PD000510; Catalase; 1.
 DR PROSITE: PS00437; CATALASE_1; 1.
 DR PROSITE: PS00438; CATALASE_2; 1.
 DR OXidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
 FT ACT_SITE 53 53
 FT ACT_SITE 126 126 BY SIMILARITY.
 FT METAL 336 336 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 479 AA; 53381 MW; EBF3CBDE6778571 CRC64;
 Query Match 93.3%; Score 42; DB 1; Length 479;
 Best Local Similarity 83.3%; Pred. No. 4.2;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWHF 6
 DB 209 WVRWHF 214
 RESULT 2
 FD6C_BRANA STANDARD; PRT; 443 AA.
 ID FD6C_BRANA
 AC F48627;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.19.-).
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seed;
 RX MEDLINE=94345008; PubMed=8066133;
 RA Hitz W.D., Carlson T.J., Booth J.R. Jr., Kinney A.J., Stecca K.L.,
 RA Yadav N.S.;
 RT "Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA
 and its expression in a cyanobacterium";
 RL Plant Physiol. 105:635-641(1994)
 CC -!- FUNCTION: Chloroplast omega-6 fatty acid desaturase introduces the
 second double bond in the biosynthesis of 16:3 and 18:3 fatty
 acids, important constituents of plant membranes. It is thought to
 use ferredoxin as an electron donor and to act on fatty acids
 esterified to galactolipids, sulfolipids and phosphatidylglycerol.
 CC -!- PATHWAY: Polyunsaturated fatty acid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
 CC -!- DOMAIN: The histidine box domains may contain the active site
 and/or be involved in metal ion binding.
 CC -!- SIMILARITY: Belongs to the fatty acid desaturase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions. There are no restrictions on its
 modified and this statement is not removed as its content is in no way
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L29214; AAA50157.1; --
 CC PIR; T08136; T08136.
 CC InterPro; IPR005804; FA desat fam.
 CC Pfam; PF00487; FA desaturase; 1.
 CC ProDom; PD001081; FA desat fam; 2.
 KW Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
 KW Transit peptide.
 FT TRANSIT 1 64 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 65 443 OMEGA-6 FATTY ACID DESATURASE.
 FT DOMAIN 166 170 HISTIDINE BOX-1.

FT DOMAIN 202 206 HISTIDINE BOX-2.
 FT DOMAIN 362 366 HISTIDINE BOX-3.
 SQ SEQUENCE 443 AA; 50755 MW; 27544B33AF2781D8 CRC64;
 Query Match 88.9%; Score 40; DB 1; Length 443;
 Best Local Similarity 85.3%; Pred. No. 8.4;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WVRWHF 6
 DB 252 WVRWHF 257
 RESULT 3
 FD6C_ARATH STANDARD; PRT; 448 AA.
 ID FD6C_ARATH
 AC P46312; Q9M094;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.19.-).
 GN FAD6 OR FADC OR AT4G30950 OR F6118.140.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE OF 1-418 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=95148736; PubMed=7846158;
 RA Falcone D.U., Gibson S., Lemieux B., Somerville C.R.;
 RT "Identification of a gene that complements an Arabidopsis mutant
 deficient in chloroplast omega 6 desaturase activity";
 RL Plant Physiol. 106:1453-1459(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller A., Stiekema W., Entian K.-D., Terryn N.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Brandt P., Grivell L., Rieger M.,
 RA Harris B., Ansong W., Brandt P., Obermaier B., Mache R., Mueller M.,
 RA Weichselgartner M., de Simone V., Puigdomenech P., Watson M., Schmidtkeini T.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtkeini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohenseil J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weljens J., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenecker T., Boche G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Berneser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Schaffe M., Grimm M., Loebner T.-H.,
 RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Quigley F., Clabaud G., Lecharny A., Aubourg S.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Cherdorf F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 RA Friskman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Farnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,

RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
 RA Nelson J., Spieth J., Ryan B., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidan M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:769-777(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv Columbia;
 EX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kaniya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tanse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 RA "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RL Science 302:842-846(2003).
 CC -!- FUNCTION: Chloroplast omega-6 fatty acid desaturase introduces
 CC the second double bond in the biosynthesis of 16:3 and 18:3 fatty
 CC acids, important constituents of plant membranes. It is thought
 CC to use ferredoxin as an electron donor and to act on fatty acids
 CC esterified to galactolipids, sulfolipids and phosphatidylglycerol.
 CC -!- PATHWAY: Polyunsaturated fatty acid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
 CC -!- DEVELOPMENTAL STAGE: Highest levels found in expanding leaves.
 CC -!- DOMAIN: The histidine box domains may contain the active site
 CC and/or be involved in metal ion binding.
 CC -!- SIMILARITY: Belongs to the fatty acid desaturase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U09503; AAA92800.1; -;
 CC EMBL; AL022198; CAA18198.1; -;
 CC EMBL; AL161578; CAB79813.1; -;
 CC EMBL; AY045621; AAK73979.1; -;
 CC EMBL; AY058078; AAL24186.1; -;
 CC EMBL; AY058852; AAL24240.1; -;
 CC PIR; D85362; D85362.
 CC InterPro: IPR005804; FA_desat_fam.
 CC Pfam; PF00487; FA_desaturase; 1.
 CC ProDom; PD001081; FA_desat_fam; 2.
 KW Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
 FT Transit peptide.
 FT TRANSIT 1 69 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 70 448 OMEGA-6 FATTY ACID DESATURASE.
 FT DOMAIN 171 175 HISTIDINE BOX-1.
 FT DOMAIN 207 211 HISTIDINE BOX-2.
 FT DOMAIN 367 371 HISTIDINE BOX-3.
 SQ SEQUENCE 448 AA; 51225 MW; C3AC72FB28F287 CRC64;
 Query Match 88.9%; Score 40; DB 1; Length 448;
 Best Local Similarity 83.3%; Pred. No. 8.5;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 Db 257 WVRWHF 262
 RESULT 4
 ZCH5 HUMAN
 ID ZCH5 HUMAN STANDARD; PRT; 475 AA.
 AC Q8N8U3;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Zinc finger, CCHC domain containing protein 5.
 GN ZCCHC5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
 RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RA "NEDO human cDNA sequencing project.";
 RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RL -!- SIMILARITY: Contains 1 CCHC-type zinc finger.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AK096184; BAC04719.1; -;
 CC Genew; HGNC:22997; ZCCHC5.
 CC InterPro: IPR001878; Znf_CCHC.
 CC Pfam; PF00098; zf_CCHC; 1.
 CC PRINTS; PR00939; C2HCZNFINGER.
 CC SMART; SM00343; ZNF_C2HC; 1.
 CC PROSITE; PS0158; ZF_CCHC; 1.
 KW Zinc-finger.
 FT DOMAIN 61 194 PRO-RICH.
 FT ZN FING 443 462 CCHC-TYPE.
 SQ SEQUENCE 475 AA; 52817 MW; 0A56AB62220F56D CRC64;
 Query Match 86.7%; Score 39; DB 1; Length 475;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db 436 WVRWH 440
 RESULT 5
 CATA ONCV
 ID CATA ONCV STANDARD; PRT; 482 AA.
 AC Q27710; O85499;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Catalase (EC 1.11.1.6).
 GN CAT.
 OS Onchocerca volvulus endobacterium.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Wolbachiae; Wolbachia.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the "phage" integrase family. Xerc
 CC subfamily 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF441790; AAM81156.1; -;
 CC HAMAP: MF_01808; -; 1.
 CC InterPro: IPR004107; Phage integr N.
 CC InterPro: IPR002104; Phage integrase.
 CC Pfam: PF02899; Phage integr N; 1.
 CC Pfam: PF00589; Phage integrase; 1.
 CC DNA recombination; DNA integration; Cell division;
 CC Chromosome partition; DNA-binding; Plasmid.
 CC ACT_SITE 167 167 BY SIMILARITY.
 CC ACT_SITE 191 191 BY SIMILARITY.
 CC ACT_SITE 264 264 BY SIMILARITY.
 CC ACT_SITE 267 267 BY SIMILARITY.
 CC ACT_SITE 291 291 BY SIMILARITY.
 CC ACT_SITE 300 300
 CC TRANSIENT COVALENT LINKAGE TO DNA DURING
 CC STRAND CLEAVAGE AND REJOINING (BY
 CC SIMILARITY).
 CC SEQUENCE 320 AA; 37022 MW; E1869F64F2AB02DC CRC64;
 CC
 CC Query Match 80.0%; Score 36; DB 1; Length 320;
 CC Best Local Similarity 80.0%; Pred. No. 28;
 CC Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 WVRWH 5
 CC : : : :
 CC Db 213 WLRWH 217
 CC
 CC RESULT 8
 CC FD6C SOYBN
 CC ID _FD6C SOYBN STANDARD; PRT; 424 AA.
 CC AC P48628;
 CC DT 01-FEB-1996 (Rel. 33, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.19.-).
 CC OS Glycine max (soybean)
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 CC OX NCBI_TaxID=3947;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC RP TISSUE=Seed;
 CC RC MEDLINE=94345008; PubMed=8066133;
 CC RA Hitz W.D.; Carlson T.J.; Booth J.R. Jr.; Kinney A.J.; Stecca K.L.;
 CC RA Yadav N.S.;
 CC RT "Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA
 CC and its expression in a cyanobacterium";
 CC RL Plant Physiol. 105:635-641(1994).
 CC -!- FUNCTION: Chloroplast omega-6 fatty acid desaturase introduces the
 CC second double bond in the biosynthesis of 16:3 and 18:3 fatty
 CC acids, important constituents of plant membranes. It is thought to
 CC use ferredoxin as an electron donor and to act on fatty acids
 CC esterified to galactolipids, sulfolipids and phosphatidylglycerol.
 CC -!- PATHWAY: Polyunsaturated fatty acid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
 CC -!- DOMAIN: The histidine box domains may contain the active site
 CC and/or be involved in metal ion binding.
 CC -!- SIMILARITY: Belongs to the fatty acid desaturase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L29215; AAA50158.1; -;
 CC PIR: T07742; T07742.
 CC InterPro: IPR005804; FA desat fam.
 CC Pfam: PF00487; FA desaturase; 1.
 CC ProDom: PD001081; FA desat fam; 2.
 CC Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
 CC Transmembrane protein.
 CC TRANSIT 1 63 CHLOROPLAST (BY SIMILARITY).
 CC CHAIN 64 424 OMEGA-6 FATTY ACID DESATURASE.
 CC DOMAIN 165 169 HISTIDINE BOX-1.
 CC DOMAIN 201 205 HISTIDINE BOX-2.
 CC DOMAIN 361 365 HISTIDINE BOX-3.
 CC SEQUENCE 424 AA; 49641 MW; 4F3DF52D4B1A2009 CRC64;
 CC
 CC Query Match 80.0%; Score 36; DB 1; Length 424;
 CC Best Local Similarity 66.7%; Pred. No. 37;
 CC Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 WVRWH 6
 CC : : : :
 CC Db 251 WLMWH 256
 CC
 CC RESULT 9
 CC BCSB XANAC
 CC ID _BCSB XANAC STANDARD; PRT; 788 AA.
 CC AC P58933;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Cyclic di-GMP binding protein precursor (Cellulose synthase regulatory
 CC subunit)
 CC GN BCSB OR XAC3517.
 CC OS Xanthomonas axonopodis (pv. citri).
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 CC OC Xanthomonadaceae; Xanthomonas.
 CC OX NCBI_TaxID=92829;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC RP STRAIN=306 / ATCC 13902 / XV 101;
 CC RC MEDLINE=22022145; PubMed=12024217;
 CC RA da Silva A.C.R.; Ferro J.A.; Reinach F.C.; Farah C.S.; Furlan L.R.;
 CC RA Quaggio R.B.; Monteiro-Vitorello C.B.; Van Sluys M.A.; Almeida N.F.;
 CC RA Alves L.M.C.; do Amaral A.M.; Bertolini M.C.; Camargo L.E.A.;
 CC RA Camarotte G.; Canavan F.; Cardozo J.; Chambergo F.; Ciapina L.P.;
 CC RA Cicarelli R.M.B.; Coutinho L.L.; Cursino-Santos J.R.; El-Dorri H.;
 CC RA Faria J.B.; Ferreira A.J.S.; Ferreira R.C.C.; Ferro M.I.T.;
 CC RA Formighieri E.F.; Franco M.C.; Greggio C.C.; Gruber A.;
 CC RA Katsuyama A.M.; Kishi L.T.; Leite R.P.; Lemos E.G.M.; Lemos M.V.F.;
 CC RA Locali E.C.; Machado M.A.; Madeira A.M.B.N.; Martinez-Rossi N.M.;
 CC RA Martins E.C.; Meidanis J.; Menck C.F.M.; Miyaki C.Y.; Moon D.H.;
 CC RA Moreira L.M.; Novo M.T.M.; Okura V.K.; Oliveira M.C.; Oliveira V.R.;
 CC RA Pereira H.A.; Rossi A.; Sena J.A.D.; Silva C.; de Souza R.F.;
 CC RA Spinola L.A.F.; Takita M.A.; Tamura R.E.; Teixeira E.C.; Tezza R.I.D.;
 CC RA Trindade dos Santos M.; Truffi D.; Tsai S.M.; White F.F.;
 CC RA Stubal J.C.; Kitajima J.P.;
 CC RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 CC host specificities";
 CC RL Nature 417:459-463(2002).
 CC -!- FUNCTION: Binds the cellulose synthase activator, bis-(3'-5')
 CC cyclic diguanylic acid (c-di-GMP) (By similarity).
 CC -!- PATHWAY: Bacterial cellulose biosynthesis.
 CC -!- SUBUNIT: Tightly associated with the cellulose synthase catalytic
 CC subunit (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE ACBS/BCSB FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AE012000; AAM38360.1; ALT INIT.
 DR InterPro; IPR003920; Cell_synth_B.
 DR Pfam; PF03170; BcsB; 1
 DR FRNTS; PRO1440; CELLSYNTHASEB.
 KW Cellulose biosynthesis; Signal; Complete proteome.
 FT SIGNAL 1 34 POTENTIAL.
 FT CHAIN 35 788 CYCLIC DI-GMP BINDING PROTEIN.
 SQ SEQUENCE 788 AA; 85169 MW; 08D7ADF415D02BBC CRC64;

Query Match 80.0%; Score 36; DB 1; Length 788;
 Best Local Similarity 66.7%; Pred. No. 69;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWF 6
 Db 743 WLRWF 748

RESULT 10

SUV3 DROME STANDARD; PRT; 1169 AA.
 AC P20193;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Suppressor of variegation protein 3-7.
 GN SUVAR(3)7.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95223788; PubMed=7708496;
 RA Cleard F., Matsarskaia M., Spierer P.;
 RT "The modifier of position-effect variegation Suvar(3)7 of Drosophila:
 RT there are two alternative transcripts and seven scattered zinc
 RT fingers, each preceded by a tryptophan box.";
 RL Nucleic Acids Res. 23:796-802(1995).
 RN [2]
 RP ERRATUM.
 RA Cleard F., Matsarskaia M., Spierer P.;
 RL Nucleic Acids Res. 23:3804-3804(1995).
 RN [3]
 RP SEQUENCE OF 255-1169 FROM N.A.
 RX MEDLINE=90190836; PubMed=2107402;
 RA Reuter G., Glarre M., Farah J., Gausz J., Spierer A., Spierer P.;
 RT "Dependence of position-effect variegation in Drosophila on dose of a
 RT gene encoding an unusual zinc-finger protein.";
 RL Nature 344:219-223(1990).
 RN [4]
 RP INTERACTION WITH SU(VAR)39.
 RX MEDLINE=21856321; PubMed=11867540;
 RA Schotta G., Ebert A., Kraus V., Fischer A., Hoffmann J., Rea S.,
 RA Jenwein T., Dorn R., Reuter G.;
 RT "Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation
 RT and heterochromatic gene silencing.";
 RL EMBO J. 21:1121-1131(2002).

CC -!- FUNCTION: This protein is a dose-limiting factor in position-
 CC effect variegation, the inactivation in some cells of a gene
 CC translocated next to heterochromatin. It could play a role in
 CC chromosome condensation.
 CC -!- SUBUNIT: Interacts with SU(VAR)39.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -----
 DR EMBL; X52187; CAA36434.1; -.
 DR PIR; S09151; S09151.
 DR FlyBase; FBgn0003598; Su(var)3-7.
 DR InterPro; IPR004210; BSS motif.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF02944; BSS; 1.
 DR SMART; SM00355; Znf_C2H2; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 KW Zinc-finger; Nuclear protein; DNA-binding; Metal-binding; Repeat.
 FT ZN_FING 136 155 C2H2-TYPE.
 FT ZN_FING 238 262 C2H2-TYPE.
 FT ZN_FING 344 365 C2H2-TYPE.
 FT ZN_FING 406 431 C2H2-TYPE.
 FT ZN_FING 526 548 C2H2-TYPE.
 FT ZN_FING 656 680 C2H2-TYPE.
 FT ZN_FING 748 771 C2H2-TYPE.
 FT DOMAIN 844 1057 BINDS TO SU(VAR)39.
 FT DOMAIN 86 92 POLY-ASP.
 FT DOMAIN 589 592 POLY-GLU.
 FT DOMAIN 780 788 POLY-ALA.
 FT DOMAIN 1008 1011 POLY-ASN.
 FT DOMAIN 1113 1116 POLY-ASN.
 SQ SEQUENCE 1169 AA; 131110 MW; A2B9380941645328 CRC64;

Query Match 80.0%; Score 36; DB 1; Length 1169;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWF 5
 Db 219 WLRWF 223

RESULT 11

YC07 METJA STANDARD; PRT; 226 AA.
 ID YC07 METJA
 AC G58604;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical acetyltransferase MJ1207 (EC 2.3.1.-).
 GN MJ1207.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kervalec A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -!- SIMILARITY: BELONGS TO THE ACETYLTRANSFERASE FAMILY. STRONG, TO
 CC A.FULGIDUS AF0521.

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67562; AAB99211.1; -.
DR PIR; F64450; F64450.
DR TIGR; MJ1207; -.
DR InterPro; IPR000182; GCS5acetyl_trans.
DR Pfam; PF00583; Acetyltransf; 1.
DR Hypothetical protein; Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 226 AA; 26939 MW; 376E18D3509E2DA CRC64;
Query Match 77.8%; Score 35; DB 1; Length 226;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWHF 6
DB 109 WARWVF 114
RESULT 12
ID HIS4 SYN7 STANDARD; PRT; 254 AA.
AC Q8GJM0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]
DE imidazole-4-carboxamide isomerase (EC 5.3.1.16)
DE (phosphoribosylformimino-5-aminoimidazole carboxamide ribotide
DE isomerase)
GN HIS4 OR SENG020.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RA Holtman C.K., Sandoval P., Chen Y., Socias T., McMurtry S.,
RA Gonzalez A., Salinas I., Golden S.S., Yoderian P.;
RA Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-5-[(5-
CC phosphoribosylamino)methylideneamino] imidazole-4-carboxamide = 5-
CC [(5-phospho-1-deoxyribulos-1-ylamino)methylideneamino]-1-(5-
CC phosphoribosyl)imidazole-4-carboxamide.
CC -1- PATHWAY: Histidine biosynthesis, fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the hisA / hisF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY157498; ANNA6174.1; -.
DR HAMAP; MF_01014; -.
DR InterPro; IPR003009; FMN enzyme.
DR InterPro; IPR006063; His4.
DR Pfam; PF00977; His_biosynth.
DR TIGRFAMs; TIGR00007; TIGR00007; 1.
DR Isomerase; Histidine biosynthesis.
SQ SEQUENCE 254 AA; 26820 MW; 16D1A158A2C45514 CRC64;
Query Match 77.8%; Score 35; DB 1; Length 254;
Best Local Similarity 80.0%; Pred. No. 32;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWH 5
DB 167 WARWH 171
RESULT 13
PLSB_CABEL
ID PLSB_CABEL STANDARD; PRT; 718 AA.
AC Q22949;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Probable glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE (EC 2.3.1.15) (GPAT).
GN F08F3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RA Blanchard M., Bradshaw H.;
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: First step in de novo phospholipid biosynthesis. It may
CC also function in the regulation of membrane biogenesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC (Potential).
CC -1- SIMILARITY: Belongs to the GPAT / DAPAT family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U64847; AAB04876.1; -.
DR PIR; T29448; T29448.
DR WormPep; F08F3.2; CE09258.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; PlsC; 1.
DR KW Phospholipid biosynthesis; Transferase; Acyltransferase;
DR Transmembrane; Mitochondrion; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 718 PROBABLE GLYCEROL-3-PHOSPHATE
FT ACYLTRANSFERASE.
FT TRANSMEM 409 425 POTENTIAL.
SQ SEQUENCE 718 AA; 82071 MW; E0A36AA86FC138D CRC64;
Query Match 77.8%; Score 35; DB 1; Length 718;
Best Local Similarity 66.7%; Pred. No. 92;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WVRWHF 6
DB 178 WCNWHF 183
RESULT 14
ID RS12 METJA STANDARD; PRT; 148 AA.
AC P54062;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S12P.
```

GN RPS12P OR MJ1046.
 OS Methanococcus jannaschii;
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8698087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kleravage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Cover B.K., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -!- FUNCTION: With S4 and S5 plays an important role in translational
 CC accuracy. Located at the interface of the 30S and 50S subunits (By
 CC similarity).
 CC -!- SUBUNIT: Part of the 30S ribosomal subunit.
 CC -!- SIMILARITY: Belongs to the S12P family of ribosomal proteins.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; U67547; AB99050.1; -;
 DR TIGR; MJ1046; -;
 DR HAMAP; MF 00403; -; 1.
 DR InterPro; IPR008994; Nucleic acid OB.
 DR InterPro; IPR006032; Ribosomal S12_23.
 DR InterPro; IPR005680; Ribosomal S23.
 DR Pfam; PF00164; Ribosomal S12; 1.
 DR ProDom; PD000576; Ribosomal_S12_23; 1.
 DR TIGRFAMs; TIGR00982; S23_S12_E_A; 1.
 DR PROSITE; PS00055; RIBOSOMAL_S12; 1.
 DR Ribosomal protein; RNA-binding; rRNA-binding; Complete proteome.
 KW RIBOSOMAL PROTEIN; RNA-BINDING; rRNA-BINDING; Complete proteome.
 SQ SEQUENCE 148 AA; 16829 MW; 6261744BDED62CFD CRC64;
 Query Match 75.6%; Score 34; DB 1; Length 148;
 Best Local Similarity 80.0%; Pred. No. 28;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db 22 WCRWH 26
 RESULT 15
 CBL_ECOLI
 ID CBL_ECOLI STANDARD; PRT; 316 AA.
 AC Q47083; P76353;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE HTH-type transcriptional regulator cbl.
 GN CBL OR B1987.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=96105196; PubMed=8529872;

RA Iwanicka-Nowicka R., Hryniewicz M.M.;
 RT "A new gene, cbl, encoding a member of the LysR family of
 RT transcriptional regulators belongs to Escherichia coli cys regulon.";
 RL Gene 166:11-17(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1232-1244(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251358; PubMed=9097040;
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshino T.,
 RA Saito N., Sampei G., Seki Y., Sivasubramanian S., Tagami H.,
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horikuchi T.;
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 40.1-50.0 min region on the linkage map.";
 RL DNA Res. 3:379-392(1996)
 CC -!- FUNCTION: MAY BE AN ACCESSORY REGULATORY PROTEIN WITHIN THE CYS
 CC REGULON.
 CC -!- SIMILARITY: Contains 1 HTH LysR-type DNA-binding domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; L31639; AAA99929.1; -;
 DR EMBL; AE000290; AAC75049.1; -;
 DR EMBL; D90837; BAA15805.1; -;
 DR PIR; C64963; C64963.
 DR HSP; P45600; IAL3.
 DR EcoGene; EG14264; cbl.
 DR InterPro; IPR000847; HTH_LysR.
 DR InterPro; IPR005119; LysR_subst.
 DR Pfam; PF00126; HTH_1; 1.
 DR Pfam; PF03466; LysR_substrate; 1.
 DR PRINTS; PR00039; HTHLYSR.
 DR PROSITE; PS50931; HTH_LYSR; 1.
 KW Transcription regulation; DNA-binding; Complete proteome.
 DOMAIN 1 59 HTH LYSR-TYPE
 FT DNA BIND 19 39 H-T-H MOTIF (POTENTIAL).
 FT CONFLICT 143 143 A -> R (IN REF. 1).
 SQ SEQUENCE 316 AA; 35856 MW; 70CB59476501067C CRC64;
 Query Match 75.6%; Score 34; DB 1; Length 316;
 Best Local Similarity 80.0%; Pred. No. 59;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db 163 WFRWH 167
 RESULT 16
 CBL_KLEAE
 ID CBL_KLEAE STANDARD; PRT; 316 AA.
 AC Q08598;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)

```

DE HTH-type transcriptional regulator cbl.
GN CBL.
OS Klebsiella aerogenes.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=28451;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W70 / KC1043;
RX MEDLINE=93209957; PubMed=8458853;
RA Schwacha A., Bender R.A.;
RT "The nac (nitrogen assimilation control) gene from Klebsiella
aerogenes.";
RL J. Bacteriol. 175:2107-2115(1993).
CC -1- FUNCTION: MAY BE AN ACCESSORY REGULATORY PROTEIN WITHIN THE CVS
REGULON.
CC -1- SIMILARITY: Contains 1 HTH LysR-type DNA-binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; L01114; AAA18174.1; -.
CC HSP; P45600; 1AL3.
CC InterPro; IPR000847; HTH_LysR.
CC InterPro; IPR005119; LysR_subst.
CC Pfam; PF00126; HTH_1; 1.
CC Pfam; PF03466; LysR_substrate; 1.
CC PRINTS; PR00039; HTHLYSR.
CC PROSITE; PS09331; HTH_LYSR; 1.
CC TRANSCRIPTION regulation; DNA-binding.
CC DOMAIN 1 59 HTH_LYSR-TYPE.
CC DNA_BIND 19 38 H-T-H MOTIF (POTENTIAL).
CC SEQUENCE 316 AA; 35702 MW; 17578BD9F50CEA7 CRC64;
Query Match 75.6%; Score 34; DB 1; Length 316;
Best Local Similarity 80.0%; Pred. No. 59;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWH 5
DB 163 WFRWH 167
RESULT 17
MTCL CAUCR STANDARD; PRT; 358 AA.
AC Q45971;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Modification methylase CcrM1 (SC 2.1.1.72) (Adenine-specific
DE methyltransferase CcrM1) (M.CcrM1).
GN CCRIM OR CCRM OR CC0378.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CB15N / NA1000;
RX MEDLINE=94118303; PubMed=8289276;
RA Zweiger G., Marczynski G., Shapiro L.;
RT "A Caulobacter DNA methyltransferase that functions only in the
predivisional cell.";
RL J. Mol. Biol. 235:472-485(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;

```

```

RX MEDLINE=21173698; PubMed=11259647;
RA Niernan W.C., Feldlyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
CC GTCM-MEDIATED METHYLATION HAS IMPORTANT CELLULAR FUNCTIONS.
CC APPEARS TO CONTRIBUTE TO THE ACCURATE CELL-CYCLE CONTROL OF DNA
REPLICATION AND CELLULAR MORPHOLOGY.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U01032; AAA18913.1; -.
CC EMBL; AB005711; AAK22365.1; -.
CC PIR; A87296; A87296.
CC PIR; S43876; S43876.
CC HSP; P11409; 1BOO.
CC REBASE; 2539; M.CcrMI.
CC TIGR; CC0378; -.
CC InterPro; IPR002295; D21N6_mtfrase.
CC InterPro; IPR001091; Met_trans_CN4.
CC InterPro; IPR002941; N6/N4_Mtase.
CC InterPro; IPR002052; N6_Mtase.
CC InterPro; IPR000051; SAM_bind.
CC Pfam; PF01555; N6_N4_Mtase; 1.
CC PRINTS; PR00506; D21N6MTFRASE.
CC PRINTS; PS00508; S21N4MTFRASE.
CC PROSITE; PS00092; N6_MTASE; 1.
CC Transferase; Methyltransferase; DNA replication; Complete proteome.
KW CONFLICT 242 242 Y -> D (IN REF. 1).
SQ SEQUENCE 358 AA; 39665 MW; 05F43266F7D4C614 CRC64;
Query Match 75.6%; Score 34; DB 1; Length 358;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WVRWH 6
DB 332 WYWHF 337
RESULT 18
GP45 HUMAN STANDARD; PRT; 372 AA.
ID GP45_HUMAN
AC Q9YSY3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable G protein-coupled receptor GPR45 (PSP24-alpha) (PSP24-1).
GN GPR45.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99156852; PubMed=10036181;
RA Marchese A., Sawzdargo M., Nguyen T., Cheng R., Heng H.H., Nowak T.,
RA Im D.-S., Lynch K.R., George S.R., O'Dowd B.P.;

```



```

CC -1- SUBCELLULAR LOCATION: Peroxisomal (Probable).
CC -1- SIMILARITY: Belongs to the flavin monooxygenase oxidase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L38858; AAA98490.1; -.
CC PIR; S5273; S5273.
CC InterPro; IPR001613; Amineoxid fl.
CC InterPro; IPR002337; Amino oxidase.
CC Pfam; PF01593; Amino oxidase; 1.
CC PRINTS; PR00757; AMINEOXIDASEF.
CC Oxidoreductase; Flavoprotein; FAD; Peroxisome.
CC NP BIND 40 95 FAD (ADP PART) (POTENTIAL).
CC SITE 493 495 MICROBODY TARGETING SIGNAL (POTENTIAL).
CC SEQUENCE 495 AA; 55616 MW; 0E614FF09D3C5B3D CRC64;
CC -----
CC Query Match 75.6%; Score 34; DB 1; Length 495;
CC Best Local Similarity 80.0%; Pred. No. 93;
CC Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC -----
CC QY 1 VWRWH 5
CC DB 94 VWRWH 98
CC -----
CC RESULT 21
CC AFSK_STRCO STANDARD; PRT; 799 AA.
CC ID AFSK_STRCO STANDARD; PRT; 807 AA.
CC AC P54741; Q9F365; Q9L002;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Serine/threonine protein kinase afsk (EC 2.7.1.37).
CC GN AFSK OR SC04423 OR SC6F11.21 OR SC6D.01.
CC OS Streptomyces coelicolor.
CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC OC Streptomycineae; Streptomycetaceae; Streptomyces.
CC OX NCBI_TaxID=1902;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=A3(2);
CC RX MEDLINE=94341568; PubMed=8063104;
CC RA Matsumoto A., Hong S.K., Ishizuka H., Horinouchi S., Beppu T.;
CC RT "Phosphorylation of the AfsK protein involved in secondary metabolism
CC in Streptomyces species by a eukaryotic-type protein kinase.";
CC RL Gene 146:47-56(1994).
CC [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=A3(2);
CC RX MEDLINE=96186909; PubMed=8635757;
CC RA Ueda K., Umeyama T., Beppu T., Horinouchi S.;
CC RT "The aerial mycelium-defective phenotype of Streptomyces griseus
CC resulting from A-factor deficiency is suppressed by a Ser/Thr kinase
CC of S. coelicolor A3(2).";
CC RL Gene 169:91-95(1996).
CC [3]
CC RP REVISIONS TO 239-240.
CC RA Matsumoto A., Hong S., Ishizuka H., Horinouchi S., Beppu T.,
CC RA Umeyama T.;
CC RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC [4]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=A3(2) / M145;
CC RX MEDLINE=21996410; PubMed=12000953;
CC RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
CC RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
CC RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

```

```

RA Cronin A., Fraser A., Coble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
RA coelicolor A3(2).";
RA Nature 417:141-147(2002).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF SECONDARY METABOLISM BY
CC PHOSPHORYLATING, ON BOTH SER AND THR, THE AFSR GLOBAL REGULATORY
CC PROTEIN INVOLVED IN THE CONTROL OF SECONDARY METABOLISM.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- PTM: Autophosphorylated on serine and threonine residues.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D45382; BAA08229.2; -.
CC EMBL; AL939140; CAD55483.1; -.
CC PhosSite; P54741; -.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00564; PQQ; 9.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.
CC KW Transferase; Serine/threonine-protein kinase; ATP-binding;
CC KW Phosphorylation; Complete proteome.
CC FT DOMAIN 16 271 PROTEIN KINASE.
CC FT NP_BIND 22 30 ATP (BY SIMILARITY).
CC FT BINDING 44 44 ATP (BY SIMILARITY).
CC FT ACT_SITE 138 138 BY SIMILARITY.
CC SEQUENCE 799 AA; 83787 MW; 4BE9BED4169F6F5B CRC64;
CC -----
CC Query Match 75.6%; Score 34; DB 1; Length 799;
CC Best Local Similarity 100.0%; Pred. No. 1.5e+02;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 2 VWRWH 6
CC DB 628 VWRWH 632
CC -----
CC RESULT 22
CC AFSK_STRGR STANDARD; PRT; 807 AA.
CC ID AFSK_STRGR STANDARD; PRT; 807 AA.
CC AC P54742;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Serine/threonine protein kinase afsk (EC 2.7.1.37).
CC GN AFSK.
CC OS Streptomyces griseus.
CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC OC Streptomycineae; Streptomycetaceae; Streptomyces.
CC OX NCBI_TaxID=1911;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=96186909; PubMed=8635757;
CC RA Ueda K., Umeyama T., Beppu T., Horinouchi S.;
CC RT "The aerial mycelium-defective phenotype of Streptomyces griseus
CC resulting from A-factor deficiency is suppressed by a Ser/Thr kinase
CC of S. coelicolor A3(2).";
CC RL Gene 169:91-95(1996).

```

CC --!- FUNCTION: INVOLVED IN THE REGULATION OF SECONDARY METABOLISM BY
CC PHOSPHORYLATING, ON BOTH SER AND THR, THE AFSR GLOBAL REGULATORY
CC PROTEIN INVOLVED IN THE CONTROL OF SECONDARY METABOLISM (BY
CC SIMILARITY).
CC --!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC --!- PTM: Autophosphorylated on serine and threonine residues (By
CC similarity).
CC --!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D45246; BAA08203.1; --
CC PhosSite; P54742; --
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_kin_AS.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00564; POQ; 4.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation.
CC FT DOMAIN 16 272 PROTEIN KINASE.
CC FT NP_BIND 22 30 ATP (BY SIMILARITY).
CC FT BINDING 44 44 ATP (BY SIMILARITY).
CC FT ACT_SITE 138 138 BY SIMILARITY.
CC SQ SEQUENCE 807 AA; 85231 MW; 66C274219155D091 CRC64;

CC Query Match 75.6%; Score 34; DB 1; Length 807;
CC Best Local Similarity 100.0%; Pred. No. 1.5e+02;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC QY 2 VRWHF 6
CC |||||
CC Db 635 VRWHF 639

CC RESULT 23
CC CUL7_HUMAN
CC ID CUL7_HUMAN STANDARD; PRT; 1698 AA.
CC AC Q14999;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Cullin homolog 7 (CUL-7).
CC GN CUL7 OR KIAA0076.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Bone marrow;
CC RX MEDLINE=96051398; PubMed=7584044;
CC RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
CC RA Seki N., Kawabayashi Y., Ishikawa K.-I., Tabata S.;
CC RT "Prediction of the coding sequences of unidentified human genes. II.
CC RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
CC RT analysis of cDNA clones from human cell line KG-1.";
CC RL DNA Res. 1:223-229(1994).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Eve;
CC RX MEDLINE=22388257; PubMed=12477932;
CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altechul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whitney M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RA [3]
RA IDENTIFICATION IN A COMPLEX WITH SKP1; FBXW8 AND RBX1.
RA MEDLINE=22388271; PubMed=12481031;
RA Dias D.C., Dolios G., Wang R., Pan Z.Q.;
RA "CUL7: A DOC domain-containing cullin selectively binds Skp1.Fbx29 to
RA form an SCF-like complex.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16601-16606(2002).
RA [4]
RA INTERACTION WITH RBX1, AND IDENTIFICATION IN A COMPLEX WITH SKP1;
RA FBXW8; RBX1 AND GLMN.
RA MEDLINE=22810107; PubMed=12904573;
RA Arai T., Kasper J.S., Skaar J.R., Ali S.H., Takahashi C.,
RA DeCaprio J.A.;
RA "Targeted disruption of p185/Cul7 gene results in abnormal vascular
RA morphogenesis";
RA Proc. Natl. Acad. Sci. U.S.A. 100:9855-9860(2003).
CC --!- FUNCTION: Component of a probable SCF-like E3 ubiquitin ligase
CC complex, which mediates the ubiquitination and subsequent
CC proteosomal degradation of target proteins. Probably plays a role
CC in the degradation of proteins involved in endothelial
CC proliferation and/or differentiation (By similarity).
CC --!- PATHWAY: Ubiquitin conjugation; third step.
CC --!- SUBUNIT: Part of a SCF-like complex consisting of CUL7, RBX1,
CC SKP1, FBXW8 and GLMN isoform 1. Interacts with a complex of SKP1
CC and FBXW8, but not with SKP1 alone.
CC --!- SIMILARITY: Belongs to the cullin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D38548; BAA07551.1; --
CC EMBL; BC033647; AAH33647.1; --
CC InterPro; IPR008938; ARM.
CC InterPro; IPR008979; Gal_bind_like.
CC PROSITE; PS01256; CULLIN_1; FALSE_NEG.
CC PROSITE; PS50069; CULLIN_2; 1.
CC XN Ubl conjugation pathway.
CC SQ SEQUENCE 1698 AA; 191188 MW; 57B11CC478E3EEDA CRC64;

CC Query Match 75.6%; Score 34; DB 1; Length 1698;
CC Best Local Similarity 80.0%; Pred. No. 3.2e+02;
CC Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC QY 1 VRWHF 5
CC |||||
CC Db 422 VRWHF 426

CC RESULT 24
CC PDRC_YEAST

ID PDRC YEAST STANDARD; PRT; 1511 AA.
 AC Q02785;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ATP-dependent permease PDR12.
 GN PDR12 OR YPL058C OR LPE14C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=97333271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albertmann K., Allen E., Ansoerge W.,
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
 RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
 RA Komp C., Kurdi O., Lashkari D., Lew H.-W., Lin A., Lin D., Louis E.J.,
 RA Marathe R., Messenguy F., Meves H.-W., Mirtipati S., Moestl D.,
 RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Puzelle D., Schafer M., Scharfe M.,
 RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
 RA Urrestazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI";
 RL Nature 387:103-105(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the ABC transporter family. PDR5 subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U39205; AAB68307.1; --
 CC F1R; S60932; S60932.
 CC Germonline; 144040; --
 CC SGD; S0005979; PDR12.
 CC GO; GO:0005886; C:plasma membrane; IDA.
 CC GO; GO:0005342; P:organic acid transporter activity; IDA.
 CC GO; GO:0015849; P:organic acid transport; IDA.
 CC GO; GO:0019541; P:protonate metabolism; ISP.
 CC InterPro; IPR003593; AAA_Arpase.
 CC InterPro; IPR003439; ABC_transporter.
 CC Pfam; PF00005; ABC_tran; 2.
 CC ProDom; PD000006; ABC_transporter; 2.
 CC SMART; SM00382; AAA; 2.
 CC PROSITE; PS00211; ABC_TRANSPORTER_1; FALSE_NEG.
 CC PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 KW ATP-binding; Transmembrane; Glycoprotein; Transport.
 FT DOMAIN 1 508
 FT TRANSMEM 509 529
 FT TRANSMEM 549 569
 FT TRANSMEM 598 618
 FT TRANSMEM 623 643
 FT TRANSMEM 658 678
 FT TRANSMEM 766 786
 FT DOMAIN 787 1182
 FT TRANSMEM 1183 1203
 FT TRANSMEM 1205 1225
 FT TRANSMEM 1255 1275
 FT TRANSMEM 1292 1312
 FT TRANSMEM 1319 1339
 FT TRANSMEM 1319 1339

FT TRANSMEM 1445 1465
 FT DOMAIN 1466 1511
 FT NP_BIND 878 885
 FT CARBOHYD 1405 1405
 SQ SEQUENCE 1511 AA; 171064 MW; 4962762AAE1997FC CRC64;
 Query Match 74.4%; Score 33.5; DB 1; Length 1511;
 Best Local Similarity 71.4%; Pred. No. 3.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 WVRW-HF 6
 Db 683 WIRWLF 689
 RESULT 25
 ID_PSD_ZYMO STANDARD; PRT; 240 AA.
 AC Q9X5E3;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE Phosphatidylserine decarboxylase proenzyme (EC 4.1.1.65) [Contains:
 DE Phosphatidylserine decarboxylase alpha chain; Phosphatidylserine
 DE decarboxylase beta chain].
 GN PSD.
 OS Zymomonas mobilis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
 OC Sphingomonadaceae; Zymomonas.
 OX NCBI_TaxID=542;
 RN [1]
 RP SEQUENCE FROM N.A. / ZM4 / CP4;
 RC STRAIN=ATCC 31821 / ZM4 / CP4;
 RA Lee H.J., Kang H.S.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Phosphatidyl-L-serine =
 CC phosphatidylethanolamine + CO(2).
 CC -1- COFACTOR: Pyruvoyl group (By similarity).
 CC -1- SIMILARITY: Belongs to the phosphatidylserine decarboxylase
 CC family. Subfamily 3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF124757; AAD29650.1; --
 CC HAMAP; MF_00664; 1
 CC InterPro; IPR003817; PS_Dcarbxyase.
 CC InterPro; IPR004428; PS_decarb_rel.
 CC Pfam; PF02666; PS_Dcarbxyase; 1.
 CC TIGRFAMs; TIGR00164; PS_decarb_rel; 1.
 CC Phospholipid biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen.
 FT CHAIN 1 203
 FT CHAIN (BY SIMILARITY).
 FT CHAIN PHOSPHATIDYL SERINE DECARBOXYLASE BETA
 FT CHAIN PHOSPHATIDYL SERINE DECARBOXYLASE ALPHA
 FT SITE 203 204
 FT MOD_RES 204 204
 FT MOD_RES CLEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).
 FT MOD_RES CONVERTED TO A PYRUVOYL GROUP (BY
 FT MOD_RES SIMILARITY).
 SQ SEQUENCE 240 AA; 26633 MW; BFA686DD7D51457 CRC64;
 Query Match 73.3%; Score 33; DB 1; Length 240;
 Best Local Similarity 80.0%; Pred. No. 66;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 WVRW-HF 6
 Db 15 IRWHF 19

RESULT 26
 YU68 CAEEL STANDARD; PRT; 244 AA.
 AC P52057;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical UPF0001 protein F09E5.8 in chromosome II.
 GN F09E5.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Chisoe S.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the UPF0001 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U37429; AAA79348.1; -;
 DR PIR; T15996; T15996.
 DR HSP; P38197; IB54.
 DR WormPep; F09E5.8; CE02615.
 DR InterPro; IPR001608; UPF0001.
 DR Pfam; PF01168; Ala racemase N; 1.
 DR TIGRFAMs; TIGR00044; TIGR00044; 1.
 DR PROSITE; PS01211; UPF0001; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 244 AA; 27195 MW; 4D6F3F3552A86A5A CRC64;
 Query Match 73.3%; Score 33; DB 1; Length 244;
 Best Local Similarity 80.0%; Pred. No. 67;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRWHP 6
 Db 77 IRWHP 81
 RESULT 27
 MTS1_RHIME STANDARD; PRT; 376 AA.
 AC O30569;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Modification methylase SmeIP (EC 2.1.1.72) (Adenine-specific
 DE methyltransferase SmeIP) (M.SmeI) (M.CcrMI).
 GN SMEIM OR CCRM OR R00926 OR SMC00021.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=97440139; PubMed=9294447;
 RA Wright R., Stephens C., Shapiro L.;
 FT "The CcrM DNA methyltransferase is widespread in the alpha subdivision
 RT of proteobacteria, and its essential functions are conserved in
 RT Rhizobium meliloti and Caulobacter crescentus."
 RL J. Bacteriol. 179:5869-5877(1997).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ramperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -!- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
 CC GATTC AND CAUSES SPECIFIC METHYLATION ON A-2 ON BOTH STRANDS.
 CC CCRM-MEDIATED METHYLATION HAS IMPORTANT CELLULAR FUNCTIONS. DNA
 CC APPEARS TO CONTRIBUTE TO THE ACCURATE CELL-CYCLE CONTROL OF DNA
 CC REPLICATION AND CELLULAR MORPHOLOGY.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
 CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF011894; AAB71350.1; -;
 DR EMBL; AL591785; CAC45498.1; -;
 DR REBASE; 3264; M.SmeIP.
 DR InterPro; IPR002295; D21N6_mtfase.
 DR InterPro; IPR001091; MetTrans_CNA.
 DR InterPro; IPR002941; N6/N4_Mtase.
 DR InterPro; IPR002052; N6_Mtase.
 DR InterPro; IPR000051; SAM bind.
 DR Pfam; PF01555; N6 N4 Mtase; 1.
 DR PRINTS; PR00506; D21N6MTFRASE.
 DR PRINTS; PR00508; S21N4MTFRASE.
 DR PROSITE; PS00092; N6_MTASE; 1.
 KW Transferase; Methyltransferase; DNA replication; Complete proteome.
 FT CONFLICT 135 141 NPMNFK -> QPDABLQ (IN REF. 1).
 FT CONFLICT 157 157 P -> A (IN REF. 1).
 SQ SEQUENCE 376 AA; 41442 MW; 790DE7FE3D22900A CRC64;
 Query Match 73.3%; Score 33; DB 1; Length 376;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VWRWHP 6

Db 347 WTFWHP 352

RESULT 28

MTB1_BRUAB STANDARD; PRT; 377 AA.

AC O30570;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Modification methylase Babi (EC 2.1.1.72) (Adenine-specific

DE methyltransferase Babi) (M.Babi) (M.CcrMI).

GN BABIM OR CCRM.

OS Brucella abortus.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Brucellaceae; Brucella.

OX NCBI_TaxID=235;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S2308;

RX MEDLINE=97440139; PubMed=9294447;

RA Wright R., Stephens C., Shapiro L.;

RT "The CcrM DNA methyltransferase is widespread in the alpha subdivision

RT of proteobacteria, and its essential functions are conserved in
 RT Rhizobium meliloti and Caulobacter crescentus.";
 RL J. Bacteriol. 179:5869-5877(1997).
 CC -!- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
 CC GANTC AND CAUSES SPECIFIC METHYLATION ON A-2 ON BOTH STRANDS.
 CC CARM-MEDIATED METHYLATION HAS IMPORTANT CELLULAR FUNCTIONS.
 CC APPEARS TO CONTRIBUTE TO THE ACCURATE CELL-CYCLE CONTROL OF DNA
 CC REPLICATION AND CELLULAR MORPHOLOGY.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
 CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF011895; AAB71351.1; -.
 DR REBASE; 3263; M.Babi.
 DR InterPro; IPR002235; D2LN6_mtfase.
 DR InterPro; IPR001091; Met_trans_CN4.
 DR InterPro; IPR002941; N6/N4_Mtase.
 DR InterPro; IPR002052; N6 Mtase.
 DR InterPro; IPR000051; SAM_bind.
 DR Pfam; PF01555; N6_N4_Mtase; 1.
 DR PRINTS; PR00506; D2LNGMTFRASE.
 DR PRINTS; PR00508; S2LN4MTFRASE.
 DR PROSITE; PS00092; N6_MTASE; 1.
 KW Transferase; Methyltransferase; DNA replication.
 SQ SEQUENCE 377 AA; 42202 MW; 657C88A25580B39D CRC64;

Query Match 73.3%; Score 33; DB 1; Length 377;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VWRWHF 6
 DB 347 WTFWHF 352

RESULT 29
 BCR_ECOLI STANDARD; PRT; 396 AA.
 AC P28246;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE Bicyclomycin resistance protein (Sulfonamide resistance protein).
 GN BCR OR BICA OR BICR OR SUR OR SUXA OR B2182.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nichols B.P.;
 RL Unpublished observations (FEB-1994).
 RN [2]
 RP SEQUENCE OF 60-396 FROM N.A.
 RC STRAIN=K12 / C600;
 RX MEDLINE=93252267; PubMed=8486276;
 RA Bentley J., Hyatt L.S., Ainley K., Parish J.H., Herbert R.B.,
 RA White G.R.;
 RT "Cloning and sequence analysis of an Escherichia coli gene conferring
 RT bicyclomycin resistance.";
 RL Gene 127:117-120(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / BHB2600;
 RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
 RA Church G.M.;

Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [5]
 RP FUNCTION
 RX MEDLINE=90146244; PubMed=2694948;
 RA Nichols B.P.; Guay G.G.;
 RT "Gene amplification contributes to sulfonamide resistance in
 RT Escherichia coli.";
 RL Antimicrob. Agents Chemother. 33:2042-2048(1989).
 CC -!- FUNCTION: Involved in sulfonamide (sulfathiazole) and
 CC bicyclomycin resistance. Probable membrane translocase.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY). BELONGS TO THE
 CC BCR/CMLA SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X63703; CAA45230.1; ALT_INIT.
 DR EMBL; U00008; AAA16406.1; ALT_INIT.
 DR EMBL; AE000308; AAC75243.1; -.
 DR PIR; E64987; E64987.
 DR EcoGene; EG11419; bcr.
 DR InterPro; IPR004734; Drug_resist.
 DR InterPro; IPR004812; Efflux_Bcr_CflA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR TIGRFAMS; TIGR00880; 2_A_01_02; 1.
 DR TIGRFAMS; TIGR00710; efflux_Bcr_CflA; 1.
 DR PROSITE; PS50850; MFS; 1.
 KW Antibiotic resistance; Transport; Transmembrane; Inner membrane;
 KW Complete proteome.
 FT TRANSMEM 9 26 POTENTIAL.
 FT TRANSMEM 49 65 POTENTIAL.
 FT TRANSMEM 77 98 POTENTIAL.
 FT TRANSMEM 106 122 POTENTIAL.
 FT TRANSMEM 138 159 POTENTIAL.
 FT TRANSMEM 166 185 POTENTIAL.
 FT TRANSMEM 215 239 POTENTIAL.
 FT TRANSMEM 251 268 POTENTIAL.
 FT TRANSMEM 286 302 POTENTIAL.
 FT TRANSMEM 309 332 POTENTIAL.
 FT TRANSMEM 345 366 POTENTIAL.
 FT TRANSMEM 373 390 POTENTIAL.
 FT CONFLICT 99 99 D -> V (IN REF. 2).
 FT CONFLICT 246 246 V -> I (IN REF. 1 AND 2).
 SQ SEQUENCE 396 AA; 43352 MW; D609AE5370E6A1D CRC64;

Query Match 73.3%; Score 33; DB 1; Length 396;
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VWRWHF 6
 DB 160 WLSWHF 165

RESULT 31

```

DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bromoperoxidase-catalase (EC 1.11.1.1).
GN BCA.
OS Streptomyces violaceus (Streptomyces venezuelae).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1936;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=ATCC 10712 / ISP5230;
RX MEDLINE=97022081; PubMed=868441;
RA Pacey S., Gross F., Vining L.C., Yang K., van Pee K.-H.;
RT "Cloning, sequencing and disruption of a bromoperoxidase-catalase
RT gene in Streptomyces venezuelae: evidence that it is not required for
RT chlorination in chloramphenicol biosynthesis.";
RL Microbiology 142:657-665 (1996).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- SIMILARITY: Belongs to the catalase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74791; CAA52796.1; -.
DR PIR; S37055; S37055.
DR HSSP; P42321; 2CAE.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
FT ACT SITE 54 54 BY SIMILARITY.
FT ACT SITE 127 127 BY SIMILARITY.
FT METAL 337 337 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 483 AA; 54087 MW; B58CF8230B8A2F55 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 483;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 6
DB 210 WVKYHF 215

RESULT 33
CATA_TOXGO
ID CATA_TOXGO STANDARD; PRT; 502 AA.
AC Q9XZD5.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Peroxisomal catalase (EC 1.11.1.6).
OS Toxoplasma gondii.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC Toxoplasma.
OX NCBI_TaxID=5811;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RH;
RX MEDLINE=20092981; PubMed=10625653;
RA Kaesch A.J., Joiner K.A.;
RT "Targeting and subcellular localization of Toxoplasma gondii
RT catalase. Identification of peroxisomes in an apicomplexan
RT parasite."
RL J. Biol. Chem. 275:1112-1118 (2000).

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=RH;
RA Ding M., Clayton C., Soldati D.;
RT "Toxoplasma gondii catalase: evidence for the existence of
RT peroxisomes in Apicomplexa.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Occurs in almost all aerobically respiring organisms and
CC serves to protect cells from the toxic effects of hydrogen
CC peroxide.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: Heme group (By similarity).
CC -!- SUBCELLULAR LOCATION: Peroxisomal.
CC -!- SIMILARITY: Belongs to the catalase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF136344; AAD30129.1; -.
DR EMBL; AF161267; AAD45528.2; -.
DR HSSP; P00432; 4BLC.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome.
FT ACT SITE 64 64 BY SIMILARITY.
FT ACT SITE 137 137 BY SIMILARITY.
FT METAL 347 347 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT SITE 500 502 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 502 AA; 57270 MW; FBA8F0551434B74D CRC64;

Query Match 73.3%; Score 33; DB 1; Length 502;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 6
DB 220 YVKWHF 225

RESULT 34
CATA_MICLU
ID CATA_MICLU STANDARD; PRT; 503 AA.
AC P29422; Q8RK91;
DT 01-APR-1993 (Rel. 25, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN KATA.
OS Micrococcus luteus (Micrococcus lysodeikticus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococccineae; Micrococccaceae; Micrococcus.
OX NCBI_TaxID=1270;
RN [1]
RP PRELIMINARY SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=93050178; PubMed=1426241;
RA Murshudov G.N., Melik-Adamyan W.R., Grebenko A.I., Barynin V.V.,
RA Vagin A.A., Vainshtein B.K., Dauter Z., Wilson K.S.;
RT "Three-dimensional structure of catalase from Micrococcus
RT lysodeikticus at 1.5-A resolution.";
RL FEBS Lett. 312:127-131 (1992).
RN [2]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (0.88 ANGSTROMS).
RX MEDLINE=22340315; PubMed=12454454;

```


RA Murshudov G.N., Grebenko A.I., Brannigan J.A., Antson A.A.,
 RA Barynin V.V., Dodson G.G., Dauter Z., Wilson K.S., Melik-Adamyan W.R.;
 RT "The structures of Micrococcus lysodeikticus catalase, its ferryl
 RT intermediate (compound II) and NADPH complex."; <http://www.isb-sib.ch/announcement/>
 RL Acta Crystallogr. D 58:1972-1982(2002).
 CC -!- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;
 CC serves to protect cells from the toxic effects of hydrogen
 CC peroxide.
 CC CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -!- COFACTOR: Heme group and NADPH.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the catalase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ438208; CAD27348.1; --
 DR PIR: S27264; S27264.
 DR PDB: 1HBZ; 16-AUG-01.
 DR PDB: 1GWE; 05-DEC-02.
 DR PDB: 1GWF; 05-DEC-02.
 DR PDB: 1GWH; 05-DEC-02.
 DR InterPro: IPR002226; Catalase.
 DR Pfam: PF00199; catalase; 1.
 DR PRINTS: PK00667; CATALASE.
 DR PRODOM: PD000510; Catalase; 1.
 DR PROSITE: PS00437; CATALASE 1; 1.
 DR PROSITE: PS00438; CATALASE 2; 1.
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide; NADP;
 KW 3D-structure.
 KW ACT_SITE 61
 FT ACT_SITE 133
 FT METAL 343
 FT METAL 343
 SQ SEQUENCE 503 AA; 56906 MW; 8C60ADEFCE0E46A09 CRC64;
 Query Match 73.3%; Score 33; DB 1; Length 503;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWHF 6
 DB 216 WKYHF 221
 RESULT 35
 SIKI YEAST
 ID SIKI YEAST STANDARD; PRT; 504 AA.
 AC Q12460;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE SIKI protein (Nucleolar protein NOP56).
 GN SIKI OR NOP56 OR YLR197W OR L8167.9.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=S288c / YPH1;
 RC MEDLINE=96040178; PubMed=7547500;
 RX Morin P.J., Downs J.A., Snodgrass A.M., Gilmore T.D.;
 RA "Genetic analysis of growth inhibition by GAL4-L kappa B-alpha in
 RT Saccharomyces cerevisiae."; <http://www.isb-sib.ch/announcement/>
 RL Cell Growth Differ. 6:789-798(1995).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=S288c / AB972;
 RX MEDLINE=97313267; PubMed=9169871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansong W.,
 RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
 RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 RA Hausen-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
 RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
 RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
 RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
 RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Urestarazu L.A., Vandenbol M., Verhasselt P.,
 RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
 RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hobeisel J.D.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 RL Nature 387:87-90(1997).
 RN [3]
 RN CHARACTERIZATION, AND MUTAGENESIS.
 RX MEDLINE=98038777; PubMed=9372940;
 RA Gautier T., Berges T., Tollervy D., Hurt E.;
 RT "Nucleolar KEX/D repeat proteins Nop56p and Nop58p interact with Nop1p
 RT and are required for ribosome biogenesis."; <http://www.isb-sib.ch/announcement/>
 RL Mol. Cell. Biol. 17:7088-7098(1997).
 CC -!- FUNCTION: Required for 60S ribosomal subunit synthesis.
 CC -!- SUBUNIT: Interacts with Nop1 and Nop58.
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
 CC -!- SIMILARITY: Belongs to the NOP5/NOP56 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U20237; AAC49066.1; --
 DR EMBL: U14913; AAB67431.1; --
 DR PIR: S48550; S48550.
 DR GenOnline: 142259; --
 DR SGD: S0004187; SIKI.
 DR GO: GO:0005732; C:small nucleolar ribonucleoprotein complex; IPI.
 DR GO: GO:0030490; P:processing of 20S pre-rRNA; IPI.
 DR InterPro: IPR002687; Nop.
 DR Pfam: PF01798; Nop; 1.
 DR ProDom: PD004104; Nop; 1.
 DR Ribosome biogenesis; Nuclear protein.
 KW DOMAIN 443 504
 FT MUTAGEN 333 333
 FT MUTAGEN 355 355
 FT MUTAGEN 385 385
 FT MUTAGEN 385 385
 FT SEQUENCE 504 AA; 56864 MW; F8522A5870EF4842 CRC64;
 Query Match 73.3%; Score 33; DB 1; Length 504;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WVRWHF 6
 DB 198 WYGHF 203
 RESULT 36
 NOP5 YEAST
 ID NOP5 YEAST STANDARD; PRT; 511 AA.
 AC Q12459;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nucleolar protein NOP58 (Nucleolar protein NOP5).

GN NOPS8 OR NOPS5 OR YOR310C OR O6108.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RA Pearson B.M., Hernandez Y., Wolf S.S., Kalogeropoulos A., Schweizer M.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=W303;
RX MEDLINE=22389165; PubMed=9632712;
RA Wu P., Brockenbrough J.S., Metcalfe A.C., Chen S., Aris J.P.;
RT "Nop5p is a small nucleolar ribonucleoprotein component required for
pre-18S rRNA processing in yeast."
RL J. Biol. Chem. 273:16453-16463(1998).
CC [1]- FUNCTION: Required for pre-18S rRNA processing. May bind
microtubules.
CC [1]- SUBUNIT: Interacts with NOP56 and NOP1.
CC [1]- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC [1]- SIMILARITY: Belongs to the NOP5/NOP56 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X90565; CAA62165.1; -.
CC EMBL; Z75217; CAA99630.1; -.
CC EMBL; AF056070; AAC39484.1; -.
CC PIR; S58322; S58322.
CC GeneOnline; 143898; -.
CC SGD; S0005837; NOP58.
CC DR GO; GO:0005732; C:small nucleolar ribonucleoprotein complex; IPI.
CC DR GO; GO:0003754; F:chaperone activity; NAS.
CC DR GO; GO:0017069; F:snRNA binding; IDA.
CC DR GO; GO:0030490; P:processing of 20S pre-rRNA; IPI.
CC DR GO; GO:0006608; P:snRNP protein-nucleus import; NAS.
CC DR InterPro; IPR002687; NOP.
CC DR Pfam; PF01798; NOP; 1.
CC DR ProDom; PD004104; NOP; 1.
CC DR Ribosome biogenesis; Nuclear
KW Ribosome biogenesis; Nuclear
FT DOMAIN 441 511 ASP/GLU/LYS-RICH.
SQ SEQUENCE 511 AA; 56956 MW; 8A2889448B2A19E2 CRC64;
Query Match 73.3%; Score 33; DB 1; Length 511;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WVRWHF 6
Db 187 WTGWHF 192
RESULT 37
NOPS HUMAN
ID NOPS_HUMAN STANDARD; PRT; 529 AA.
AC Q9Y2X3; Q9P036; Q9UFN3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nucleolar protein NOP5 (Nucleolar protein 5) (NOP58) (HSPC120).
GN NOPS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RA Lyman S.K., Gerace L.;
RT "Cloning and characterization of NOP5/NOP58."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20384617; PubMed=10925205;
RA Nelson S.A., Santora K.E., LaRoche W.J.;
RT "Isolation and characterization of a novel PDGF-induced human gene."
RL Gene 253:87-93(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bonfield G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 1-447 FROM N.A.
RC TISSUE=Brain;
RA Bloeker H., Boeher M., Brandt P., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 64-529 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.,
RT "Cloning and functional analysis of cDNAs with open reading frames for
300 previously undefined genes expressed in CD34+ hematopoietic
stem/progenitor cells".
RT Genome Res. 10:1546-1560(2000).
CC [1]- FUNCTION: Required for 60S ribosomal subunit biogenesis (By
similarity).
CC [1]- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC [1]- TISSUE SPECIFICITY: Ubiquitous.
CC [1]- SIMILARITY: Belongs to the NOP5/NOP56 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF123534; AAD27610.1; -.
CC EMBL; AF263608; AAF91394.1; -.
CC EMBL; BC032592; AAH32592.1; -.
CC EMBL; AL117554; CAB55989.1; -.
CC EMBL; AF161469; AAF23084.1; -.
CC PIR; T17299; T17299.
CC SWISS-2DPAGE; Q9Y2X3; HUMAN.
CC GO; GO:0005730; C:nucleolus; TAS.

DR GO:0003730; C:nucleolus; 133.
DR GO:0003754; F:chaperone activity; TAS.
DR GO:0030519; F:snRNP binding; TAS.

```
FT TRANSMEM 504 524 POTENTIAL.
FT TRANSMEM 542 562 POTENTIAL.
FT TRANSMEM 575 595 POTENTIAL.
SQ SEQUENCE 614 AA; 67588 MW; 669948720D4F2D2 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 614;
Best Local Similarity 66.7%; Pred.No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 WVRWHF 6
Db 167 WVRWSY 172

RESULT 40
TRKB MOUSE
ID TRKB MOUSE STANDARD; PRT; 821 AA.
AC P15209;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE BDNF/NT-3 growth factor precursor (EC 2.7.1.112) (TrkB
DE tyrosine kinase) (GP145-TrkB/GP95-TrkB) (TrkB-B).
GN NTRK2 OR TRKB
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI TaxID=10090;
RN [1] _SEQUENCE FROM N.A. (ISOFORM GP145-TRKB).
RP SEQUENCE FROM N.A. (ISOFORM GP145-TRKB).
RC TISSUE=Brain;
RX MEDLINE=90059970; PubMed=2555172;
RA Klein R., Parada L.F., Coulter F., Barbacid M.;
RT "trkB, a novel tyrosine protein kinase receptor expressed during
RT mouse neural development.";
RL EMBO J. 8:3701-3709(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS GP145-TRKB AND GP95-TRKB).
RC TISSUE=Brain;
RX MEDLINE=90263089; PubMed=2160854;
RA Klein R., Conway D., Parada L.F., Barbacid M.;
RT "The trkB tyrosine protein kinase gene codes for a second neurogenic
RT receptor that lacks the catalytic kinase domain.";
RL Cell 61:647-656(1990).
RN [3]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS L1 AND L10).
RC TISSUE=Trigeminal ganglion;
RX MEDLINE=97294706; PubMed=9148911;
RA Ninkina N., Grashchuck M., Buchman V.L., Davies A.M.;
RT "trkB variants with deletions in the leucine-rich motifs of the
RT extracellular domain.";
RL J. Biol. Chem. 272:13019-13025(1997).
RN [4]
RP FUNCTION.
RX MEDLINE=91249395; PubMed=1645620;
RA Soppet D., Escandon E., Margos J., Middlemas D.S., Reid S.W.,
RA Blair J., Burton L.E., Stanton B.R., Kaplan D.R., Hunter T.,
RA Nicolics K., Parada L.F.;
RT "The neurotrophic factors brain-derived neurotrophic factor and
RT neurotrophin-3 are ligands for the trkB tyrosine kinase receptor.";
RL Cell 65:895-903(1991).
CC -1- FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),
CC NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT NERVE GROWTH FACTOR
CC (NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE
CC NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN
CC SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-
CC GAMMA-1.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
```

```
CC Event-Alternative splicing; Named isoforms=4;
CC Comment-Additional isoforms seem to exist;
CC Name=GP145-TRKB; Synonyms=L3;
CC IsoId=P15209-1; Sequence=Displayed;
CC Name=GP95-TRKB; Synonyms=T1;
CC IsoId=P15209-2; Sequence=VSP_002908, VSP_002909;
CC Name=L1;
CC IsoId=P15209-3; Sequence=VSP_002907;
CC Name=L10;
CC IsoId=P15209-4; Sequence=VSP_002905, VSP_002906;
CC TISSUE SPECIFICITY: THE DIFFERENT FORMS ARE DIFFERENTIALLY
CC EXPRESSED IN VARIOUS CELL TYPES.
CC -1- PTM: Ligand-mediated auto-phosphorylation.
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.
CC -1- SIMILARITY: Contains 2 leucine-rich (LRR) repeats.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33385; AAA0482.1; -
CC EMBL; X17647; CAA35636.1; -
CC PIR; S06943; S06943.
CC HSP; P06213; IIRK.
CC MGD; MGI:97384; Ntrk2.
CC GO; GO:0005829; C:cytosol; IDA.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_C2.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000483; LRR Cterm.
CC InterPro; IPR000372; LRR Nterm.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR002011; RecepttyrkineII.
CC InterPro; IPR001245; Tyr_pkinase.
CC InterPro; IPR008266; Tyr_pkinase_AS.
CC Pfam; PF00047; IG; 1.
CC Pfam; PF00560; LRR; 1.
CC Pfam; PF01463; LRRCT; 1.
CC Pfam; PF01462; LRRNT; 1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot kinase; 1.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00082; LRRCT; 1.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00219; TyrKc; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00109; PROTEIN KINASE Tvr; 1.
CC PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
CC Transferase; Tyrosine-protein kinase; Transmembrane; ATP-binding;
CC Phosphorylation; Receptor; Glycoprotein; Neurogenesis; Signal;
CC Leucine-rich repeat; Repeat; Immunoglobulin domain;
CC Alternative splicing.
CC SIGNAL 1 31
CC CHAIN 32 821 BDNF/NT-3 GROWTH FACTORS RECEPTOR.
CC DOMAIN 32 429 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 430 433 POTENTIAL.
CC DOMAIN 454 821 CYTOPLASMIC (POTENTIAL).
CC REPEAT 72 93 LRR 1.
CC REPEAT 96 117 LRR 2.
CC DOMAIN 197 282 IG-LIKE C2-TYPE 1.
CC DOMAIN 301 365 IG-LIKE C2-TYPE 2.
CC DOMAIN 537 806 PROTEIN KINASE.
CC NP BIND 543 551 ATP (BY SIMILARITY).
CC BINDING 571 571 ATP (BY SIMILARITY).
```

```

FT ACT SITE 675 675 BY SIMILARITY.
FT DISULFID 32 38 BY SIMILARITY.
FT DISULFID 36 45 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 154 194 BY SIMILARITY.
FT DISULFID 218 266 BY SIMILARITY.
FT DISULFID 302 345 BY SIMILARITY.
FT MOD_RES 515 515 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 701 701 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 705 705 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 706 706 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 816 816 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 515 515 INTERACTION WITH SHC PROTEIN
FT SITE (BY SIMILARITY).
FT SITE 816 816 INTERACTION WITH PLC-GAMMA-1
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPATIC 71 71 I -> M (in isoform L10).
FT VARSPATIC 72 143 Missing (in isoform L10).
FT VARSPATIC 72 120 /FTid=VSP_002906.
FT VARSPATIC 466 476 Missing (in isoform L1).
FT VARSPATIC 477 821 PAVSINSDSDS -> FVLPHKIPLDG (in isoform GP95-TRKB).
FT VARSPATIC 821 AA; 92133 MW; 50508D5FF86D8F30 CRC64;
FT VARSPATIC 73.3%; Score 33; DB 1; Length 821;
FT VARSPATIC 70.0%; Pred. No. 2.3e+02;
FT VARSPATIC 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db 4 WLKWH 8

```

RESULT 41

```

TRKB_RAT ID TRKB_RAT STANDARD; PRT; 821 AA.
AC Q63604; Q63605; Q63606;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE BDNF/NF-3 growth factors receptor precursor (EC 2.7.1.112) (TrkB
DE tyrosine kinase) (GP145-TRKB/GP95-TRKB) (TrkB-B).
GN NTRK2 OR TRKB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Cerebellum;
RX MEDLINE=91094826; PubMed=1846020;
RA Middlemas D.S., Lindberg R.A., Hunter T.;
RT "trkb, a neural receptor protein-tyrosine kinase: evidence for a
RT full-length and two truncated receptors."
RN Mol. Cell. Biol. 11:143-153(1991).
RN [2]
RN PHOSPHORYLATION SITES.

```

```

RX MEDLINE=94149017; PubMed=8106527;
RA Middlemas D.S., Meisenhelder J., Hunter T.;
RT "Identification of TrkB autophosphorylation sites and evidence that
RT phospholipase C-gamma 1 is a substrate of the TrkB receptor."
RL J. Biol. Chem. 269:5458-5466(1994).
CC -I- FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),
CC NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT NERVE GROWTH FACTOR
CC (NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE
CC NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN
CC SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-
CC GAMMA-1.
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -I- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=GP145-TRKB; Sequence=Displayed;
CC IsoId=Q63604-1; Sequence=GP95-TRKB;
CC Name=T1; Synonyms=GP95-TRKB;
CC IsoId=Q63604-2; Sequence=VSP_002910, VSP_002911;
CC Name=T2;
CC IsoId=Q63604-3; Sequence=VSP_002912, VSP_002913;
CC -I- TISSUE SPECIFICITY: THE DIFFERENT FORMS ARE DIFFERENTIALLY
CC EXPRESSED IN VARIOUS CELL TYPES. THE T2 ISOFORM IS PRIMARILY
CC EXPRESSED IN NEURONS.
CC -I- PTM: Ligand-mediated auto-phosphorylation.
CC -I- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.
CC -I- SIMILARITY: Contains 2 leucine-rich (LRR) repeats.
CC -I- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; M55291; AAA42279.1; -
CC EMBL; M55292; AAA42280.1; -
CC EMBL; M55293; AAA42281.1; -
CC PIR; A39667; A39667.
CC PIR; B39667; B39667.
CC PIR; C39667; C39667.
CC HSSP; P06213; LTRK.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_c2.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000483; LRR_Cterm.
CC InterPro; IPR000372; LRR_Nterm.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002011; RecepttyrkinII.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR008266; Tyr_kinase_AS.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00560; LRR; 1.
CC Pfam; PF01463; LRRCT; 1.
CC Pfam; PF01462; LRRNT; 1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00082; LRRCT; 1.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS50835; IG LIKE; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
KW Transferase; Tyrosine-protein kinase; Transmembrane; ATP-binding;
KW Phosphorylation; Receptor; Glycoprotein; Neurogenesis; Signal;
KW Leucine-rich repeat; Repeat; Immunoglobulin domain;
KW Alternative splicing.
FT SIGNAL 1 31
FT CHAIN 32 821
FT DOMAIN 32 429
FT TRANSMEM 430 453
FT DOMAIN 454 821
FT REPEAT 72 93
FT REPEAT 96 117
FT DOMAIN 197 282
FT DOMAIN 295 365
FT DOMAIN 537 806
FT NP_BIND 543 551
FT BINDING 571 571
FT ACT_SITE 675 675
FT DISULFID 32 38
FT DISULFID 36 45
FT DISULFID 152 176
FT DISULFID 154 194
FT DISULFID 218 266
FT DISULFID 302 345
FT MOD_RES 515 515
FT MOD_RES 701 701
FT MOD_RES 705 705
FT MOD_RES 706 706
FT MOD_RES 816 816
FT SITE 515 515
FT SITE 816 816
FT CARBOHYD 67 67
FT CARBOHYD 95 95
FT CARBOHYD 121 121
FT CARBOHYD 178 178
FT CARBOHYD 205 205
FT CARBOHYD 241 241
FT CARBOHYD 254 254
FT CARBOHYD 280 280
FT CARBOHYD 325 325
FT CARBOHYD 338 338
FT CARBOHYD 411 411
FT VARSPPLIC 466 466
FT VARSPPLIC 477 821
FT VARSPPLIC 466 474
FT VARSPPLIC 475 821
SQ SEQUENCE 821 AA; 92186 MW; 0DDACDA212CDAA0E CRC64;
Query Match 73.3%; Score 33; DB 1; Length 821;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VWRWH 5
Db 4 WPRWH 8
RESULT 42
FDNG ECOLI
ID FDNG ECOLI STANDARD; PRT; 1015 AA.
AC P241B3; P78261;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formate dehydrogenase, nitrate-inducible, major subunit (EC 1.2.1.2)
DE (Formate dehydrogenase-N alpha subunit) (FDH-N alpha subunit)
(Anaerobic formate dehydrogenase major subunit).
FMG OR B1474.
Escherichia coli.
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
NCBI_TaxID=562;
[1]
SEQUENCE FROM N.A.
STRAIN=K12; PubMed=1834669;
MEDLINE=92042178;
Berg B.L., Li J., Heider J., Stewart V.;
"Nitrate-inducible formate dehydrogenase in Escherichia coli K-12. I.
Nucleotide sequence of the fdnGHI operon and evidence that opal (UGA)
encodes selenocysteine."
J. Biol. Chem. 266:22380-22385 (1991).
[2]
SEQUENCE FROM N.A.
STRAIN=K12 / MG1655;
MEDLINE=97426617; PubMed=9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
"The complete genome sequence of Escherichia coli K-12."
Science 277:1453-1474 (1997).
[3]
SEQUENCE FROM N.A.
STRAIN=K12; PubMed=9097039;
MEDLINE=97251357;
Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Iseno K., Itoh T.,
Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
Sampei G., Seki Y., Sivasubramanian S., Tagami H., Takeda J.,
Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
"A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map."
DNA Res. 3:363-377 (1996).
[4]
X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
MEDLINE=21882259; PubMed=11884747;
Jermakka M., Tornroth S., Byrne B., Iwata S.;
"Molecular basis of proton motive force generation: structure of
formate dehydrogenase-N."
Science 295:1863-1868 (2002).
-!- FUNCTION: Formate dehydrogenase allows E.coli to use formate as
major electron donor during anaerobic respiration, when nitrate is
used as electron acceptor. The alpha subunit forms the active
site.
-!- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
-!- COFACTOR: Binds 1 molybdenum ion, 2 molybdopterin guanine
dinucleotide (MGD) groups, 1 4Fe-4S cluster, and 1 selenocysteine
per subunit. The active-site selenocysteine is encoded by the opal
codon, UGA.
-!- PATHWAY: Anaerobic nitrate respiration.
-!- SUBUNIT: Trimer of heterotrimers, consisting of subunits alpha,
beta and gamma.
-!- SUBCELLULAR LOCATION: Periplasmic.
-!- INDUCTION: By nitrate under anaerobic conditions.
-!- SIMILARITY: Belongs to the prokaryotic molybdopterin-containing
oxidoreductase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M75029; -; NOT ANNOTATED_CDS.
EMBL; AE000244; AAD13438.1; -.
EMBL; D90788; BAA15123.1; -.

```

DR EMBL; D90789; BAA15132.1; -.
DR PIR; E64900; JS0628.
DR PDB; 1KQF; 15-MAR-02.
DR PDB; 1KQG; 15-MAR-02.
DR EcoGene; EGI1227; fdnG.
DR InterPro; IPR009010; Asp_decarb_fold.
DR InterPro; IPR006443; Formate-dh_alph.
DR InterPro; IPR006657; Mol_dinuc_bind.
DR InterPro; IPR006963; Molybdop_Fe4S4.
DR InterPro; IPR006656; Molybdopterin.
DR InterPro; IPR006655; Prok_Mboxred.
DR InterPro; IPR006311; Tat_Fe4S4; 1.
DR Pfam; PF04879; Molybdop_Fe4S4; 1.
DR Pfam; PF0384; molybdopterin; 1.
DR Pfam; PF01568; Molybdop binding; 1.
DR TIGRFAMs; TIGR01553; formate-DH-alph; 1.
DR TIGRFAMs; TIGR01409; Tat signal seg; 1.
DR PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; 1.
DR PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; FALSE_NEG.
DR PROSITE; PS00932; MOLYBDOPTERIN_PROK_3; 1.
KW Oxidoreductase; Molybdenum; Selenocysteine; Selenium; NAD;
KW Iron-sulfur; 4Fe-4S; Periplasmic; 3D-structure; Complete proteome.
FT METAL 50 50
FT METAL 53 53
FT METAL 57 57
FT METAL 92 92
FT METAL 196 196
FT SE CYS 196 196
FT CONFLICT 96 96
FT CONFLICT 484 491
FT CONFLICT 941 941
FT TURN 39 42
FT STRAND 44 49
FT STRAND 53 53
FT TURN 54 55
FT STRAND 58 64
FT TURN 69 70
FT STRAND 74 80
FT TURN 82 83
FT TURN 85 89
FT HELIX 93 96
FT TURN 97 98
FT HELIX 99 102
FT TURN 103 103
FT TURN 105 106
FT STRAND 107 107
FT STRAND 112 114
FT TURN 116 117
FT STRAND 122 123
FT HELIX 126 144
FT STRAND 146 148
FT TURN 150 151
FT STRAND 154 158
FT STRAND 161 164
FT TURN 167 168
FT HELIX 171 183
FT TURN 184 185
FT STRAND 186 186
FT STRAND 189 190
FT HELIX 192 195
FT TURN 196 197
FT HELIX 198 208
FT TURN 217 217
FT HELIX 218 222
FT STRAND 225 229
FT HELIX 233 236
FT TURN 238 241
FT HELIX 242 250
FT STRAND 254 258
FT HELIX 264 268
FT TURN 268 271
FT STRAND 271 273
FT TURN 277 278

FT HELIX 293
FT TURN 294
FT HELIX 299
FT STRAND 306
FT TURN 308
FT STRAND 310
FT TURN 314
FT STRAND 318
FT TURN 320
FT STRAND 322
FT TURN 324
FT STRAND 326
FT TURN 328
FT STRAND 332
FT TURN 336
FT STRAND 340
FT STRAND 342
FT TURN 344
FT STRAND 348
FT STRAND 350
FT TURN 352
FT TURN 357
FT STRAND 359
FT HELIX 360
FT TURN 369
FT HELIX 373
FT HELIX 384
FT HELIX 396
FT TURN 400
FT STRAND 404
FT HELIX 410
FT TURN 416
FT HELIX 418
FT TURN 433
FT TURN 435
FT TURN 439
FT STRAND 442
FT STRAND 448
FT TURN 449
FT HELIX 451
FT TURN 457
FT TURN 461
FT STRAND 463
FT HELIX 465
FT STRAND 468
FT STRAND 472
FT TURN 474

Query Match 73.3%; Score 33; DB 1; Length 1015;

Best Local Similarity 50.0%; Pred. No. 2.8e+02;

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVRWHF 6

Db 632 WLQWHW 637

RESULT 43

FDOG_ECOLI STANDARD; PRT; 1016 AA.

AC P32176; P78131;

DT 01-OCT-1993 (Rel. 27, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Formate dehydrogenase-O, major subunit (EC 1.2.1.2) (Formate

DE dehydrogenase-O alpha subunit) (PDH-Z alpha subunit) (Aerobic formate

DE dehydrogenase major subunit).

GN FDOG OR B3894.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RA Plunkett G. III, Burland V.D., Daniels D.L., Plattner F.R.;
 RT "Analysis of the *Escherichia coli* genome. III. DNA sequence of the
 RL Nucleic Acids Res. 21:3391-3398(1993).
 RN [2]
 RP REVISIONS TO 252-261; 344-348 AND 822.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE OF 1-190 FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=K12;
 RX MEDLINE=96099298; PubMed=8522521;
 RA Abaibou H., Pommer J., Giordano G., Mandrand-Berthelot M.-A.;
 RT "Expression and characterization of the *Escherichia coli* *fdo* locus
 RT and a possible physiological role for aerobic formate
 RT dehydrogenase.";
 RL J. Bacteriol. 177:7141-7149(1995).
 CC -!- FUNCTION: Allows to use formate as major electron donor during
 CC aerobic respiration. Subunit alpha possibly forms the active
 CC site.
 CC -!- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -!- COFACTOR: Molybdenum (molybdopterin) and selenocysteine. The
 CC active-site selenocysteine is encoded by the *opd* codon, UGA.
 CC May bind a 4Fe-4S cluster.
 CC -!- SUBUNIT: FORMATE DEHYDROGENASE IS A MEMBRANE-BOUND COMPLEX, FORMED
 CC BY SUBUNITS ALPHA, BETA AND GAMMA.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: Belongs to the prokaryotic molybdopterin-containing
 CC oxidoreductase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L19201; AAB03027.2; -;
 DR EMBL; AE000464; RAD13456.1; ALT_SEQ.
 DR EMBL; X87583; CAA60887.1; -;
 DR PIR; A65195; S40838.
 DR HSSP; P07658; 1FDO.
 DR Ecogene; EGI1858; fdoG.
 DR InterPro; IPR009010; Asp decarb fold.
 DR InterPro; IPR006443; Formate-dh-alpha.
 DR InterPro; IPR006657; Mol dinuc bind.
 DR InterPro; IPR006963; Molybdop Fe4S4.
 DR InterPro; IPR006656; MolybdopFe4S4.
 DR InterPro; IPR006655; MolybdopFe4S4.
 DR InterPro; IPR006311; Tat.
 DR Pfam; PF04879; Molybdop Fe4S4; 1.
 DR Pfam; PF00384; molybdopFe4S4; 1.
 DR Pfam; PF01568; molybdop binding; 1.
 DR TIGRFAMs; TIGR01553; formate-DH-alpha; 1.
 DR TIGRFAMs; TIGR01409; Tat signal seq; 1.
 DR PROSITE; PS00551; MolybdopFe4S4; 1.
 DR PROSITE; PS00490; MolybdopFe4S4; 1.
 DR PROSITE; PS00932; MolybdopFe4S4; 1.
 DR PROSITE; PS00932; MolybdopFe4S4; 1.
 KW Oxidoreductase; Molybdenum; Selenocysteine; Selenium; NAD;
 IRON-sulfur; 4Fe-4S; Complete proteome.
 FT METAL 50 50 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 53 53 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 57 57 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 92 92 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT SE_CYS 196 196 GAKLIVDPD -> RREADCDRSC (IN REF. 1).
 FT CONFLICT 252 261 ENGFA -> GKRLR (IN REF. 1).
 FT CONFLICT 344 348
 SQ SEQUENCE 1016 AA; 112502 MW; 95C06BD9633C0A7C CRC64;
 Query Match 73.3%; Score 33; DB 1; Length 1016;
 Best Local Similarity 50.0%; Pred. No. 2.8e+02;
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 WVRWHF 6
 Db 632 WLQWHW 637
 RESULT 44
 ID CRGB RALSO STANDARD; PRT; 126 AA.
 AC Q8XZK2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Protein crcb homolog.
 GN CRGB OR RSC1333 OR RS02855.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OC NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brotter P., Camus J.C., Cattolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gagin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";
 RL Nature 415:497-502(2002).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the crcb family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL646064; CAD15035.1; -;
 DR HAMAP; MF 00454; -;
 DR InterPro; IPR003691; Camphor_CrcB.
 DR Pfam; PF02537; CRCB; 1.
 DR TIGRFAMs; TIGR00494; crcb; 1.
 KW Transmembrane; Complete proteome.
 FT TRANSMEM 5 24 POTENTIAL.
 FT TRANSMEM 34 56 POTENTIAL.
 FT TRANSMEM 68 90 POTENTIAL.
 FT TRANSMEM 100 122 POTENTIAL.
 SQ SEQUENCE 126 AA; 13101 MW; C2443FBAE5C81CB3 CRC64;
 Query Match 71.1%; Score 32; DB 1; Length 126;
 Best Local Similarity 66.7%; Pred. No. 50;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 WVRWHF 6
 Db 18 WLRWAF 23
 RESULT 45


```

YH2 YEAST
ID YH2 YEAST STANDARD; PRT; 129 AA.
AC P38857;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 14.6 kDa protein in RRC104-SOL3 intergenic region.
GN YHR162W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9288c / AB972;
RC MEDLINE=94378003; PubMed=8091229;
RX Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L.W., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevasakis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL Science 265:2077-2082(1994).
CC -!- SIMILARITY: BELONGS TO THE UPF0041 (0-44) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U00027; AAB68009.1; -.
CC PIR; S48902; S48902.
CC Germonline; 139480; -.
CC SGD; S0001205; YHR162W.
CC InterPro; IPR005336; UPF0041.
CC Pfam; PF03650; UPF0041; 1.
CC Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 129 AA; 14555 MW; 5C1FB3A463DE24A7 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 129;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VWRWHF 6
| | | |
Db 64 WTRWSF 69

RESULT 46
ID YG56 YEAST STANDARD; PRT; 146 AA.
AC P53311;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 16.2 kDa protein in PFK1-TDS4 intergenic region.
GN YGR243W OR G8620.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RC MEDLINE=97245298; PubMed=9090057;
RX Guerreiro P., Azevedo D., Barreiros T., Rodrigues-Pousada C.;
RA "Sequencing of a 9.9 kb segment on the right arm of yeast chromosome
RT

```

```

RT VII reveals four open reading frames, including PFK1, the gene coding
RT for succinyl-CoA synthetase (beta-chain) and two ORFs sharing
RT homology with ORFs of the yeast chromosome VIII."
RL Yeast 13:275-280(1997).
CC -!- SIMILARITY: BELONGS TO THE UPF0041 (0-44) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z73028; CAA97272.1; -.
CC PIR; S64569; S64569.
CC Germonline; 141585; -.
CC SGD; S0003475; YGR243W.
CC InterPro; IPR005336; UPF0041.
CC Pfam; PF03650; UPF0041; 1.
CC Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 146 AA; 16230 MW; E0B13933AB142B4E CRC64;

Query Match 71.1%; Score 32; DB 1; Length 146;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VWRWHF 6
| | | |
Db 64 WTRWSF 69

RESULT 47
ID LIPB THETN STANDARD; PRT; 228 AA.
AC Q8R9EO;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lipoate-protein ligase B (EC 6.-.-) (Lipoate biosynthesis protein
DE B).
DE DE B).
DE LIPB OR TTE1673.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RC MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of T. tengcongensis genome."
RL Genome Res. 12:689-700(2002).
CC -!- FUNCTION: Involved in the attachment of lipoyl groups to proteins,
CC by creating an amide linkage that joins the free carboxyl group of
CC lipoyl acid to the epsilon-amino group of a specific lysine
CC residue in lipoylated proteins (By similarity).
CC -!- PATHWAY: Lipoate biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the lipB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE013121; AAM24874.1; -.
CC HAMAP; MF_00013; -.

```

```
DR InterPro; IPR004143; BPL_LipA_LipB.
DR InterPro; IPR000544; Lipocate B.
DR Pfam; PF03099; BPL_LipA_LipB; 1.
DR ProDom; PD006086; Lipocate B; 1.
DR TIGRFAMs; TIGR00214; lipB; 1.
DR PROSITE; PS01313; LipB; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 228 AA; 25868 MW; EF5B2722BF20F32A CRC64;

Query Match 71.1%; Score 32; DB 1; Length 228;
Best Local Similarity 60.0%; Pred. No. 91;
Matches 3; Conservative 1; Mismatches 0; Gaps 0;

QY 1 WVRWH 5
DB 151 WITHH 155

RESULT 48
COBS_SYNY3
ID COBS_SYNY3 STANDARD; PRT; 260 AA.
AC Q55714;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cobalamin synthase (EC 2.1.1.1).
GN COBS OR SLR0636.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Saruka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res 3:109-136(1996)
CC -1- FUNCTION: Joins Ado-cobinamide-GDP and alpha-ribazole to generate
CC adenosylcobalamin (Ado-cobalamin) (by similarity).
CC -1- CATALYTIC ACTIVITY: GDP-cobinamide + alpha-ribazole = cobalamin +
CC GMP.
CC -1- PATHWAY: Cobalamin biosynthesis; last step.
CC -1- SIMILARITY: Belongs to the cobs family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D64002; BRA10355.1; -.
CC PIR; S76509; S76509.
CC HANAP; MF_00719; -. 1.
DR InterPro; IPR003805; Cobs_synth.
DR Pfam; PF02654; Cobs; 1.
DR TIGRFAMs; TIGR00317; cobs; 1.
KW Cobalamin biosynthesis; Transferase; Complete proteome.

SQ SEQUENCE 260 AA; 28567 MW; 7B47AC3CDC8C3107 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 260;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 6
DB 4 WRRWF 9

RESULT 49
YAH_ECOLI
ID YAH_ECOLI STANDARD; PRT; 287 AA.
AC P77297;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yahE.
GN YAH OR B0319.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KL2 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RX Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RX Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RX Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000139; AAC73422.1; -.
CC PIR; G64758; G64758.
CC ECoGene; EG13589; yahE.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 287 AA; 32265 MW; 9EC90F5F93DF5CC CRC64;

Query Match 71.1%; Score 32; DB 1; Length 287;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 109 WLGWH 113

RESULT 50
YJLA_BACSU
ID YJLA_BACSU STANDARD; PRT; 324 AA.
AC O34428;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein yJLA.
```

GN YJLA OR BSU12260.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Rivotta C., Soldo B., Lazarevic V., Joris B., Mauel C., Karamata D.;
 RT "A 35.7 kb DNA fragment from Bacillus subtilis chromosome containing a
 RT putative 12.3 kb operon involved in hexuronate catabolism and a
 RT perfect catabolite-responsive element."
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis."
 RL Nature 390:249-256(1997).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 CC or send an email to license@sib-sib.ch).

CC EMBL; AF015825; AAC46322.1; --
 DR EMBL; Z991110; CAB13083.1; --
 DR PIR; G69851; G69851.
 DR Subtilist; BG13200; yJLA.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 5 24 Potential.
 FT TRANSMEM 39 61 Potential.
 FT TRANSMEM 68 90 Potential.
 FT TRANSMEM 95 117 Potential.
 FT TRANSMEM 130 152 Potential.
 FT TRANSMEM 162 179 Potential.
 FT TRANSMEM 199 218 Potential.
 FT TRANSMEM 228 250 Potential.
 SQ SEQUENCE 324 AA; 35624 MW; AID28B9AFB9CFA95 CRC64;

Query Match

71.1%; Score 32; DB 1; Length 324;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 WVRWHF 6
 |::|
 DB 70 WIKWSF 75

Search completed: June 10, 2004, 10:48:34
 Job time : 15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:43:04 ; Search time 33 Seconds
(without alignments)

57.367 Million cell updates/sec

Title: US-09-912-414-2

Perfect score: 45

Sequence: 1 WVRWHP 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

SPTREMBL 25:*

1: sp archaea:*

2: sp bacteria:*

3: sp fungi:*

4: sp human:*

5: sp invertebrate:*

6: sp mammal:*

7: sp mhc:*

8: sp organelle:*

9: sp phage:*

10: sp plant:*

11: sp rodent:*

12: sp virus:*

13: sp vertebrate:*

14: sp unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	42	93.3	479	16	Q88QK9	Q88QK9 pseudomonas
2	41	91.1	256	10	Q9FIN2	Q9FIN2 arabidopsis
3	41	91.1	325	10	Q8LD77	Q8LD77 arabidopsis
4	41	91.1	325	10	Q84WW2	Q84WW2 arabidopsis
5	40	88.9	201	10	Q9C600	Q9C600 arabidopsis
6	40	88.9	202	10	Q9AXY3	Q9AXY3 brassica ca
7	40	88.9	213	10	Q9AXY5	Q9AXY5 brassica na
8	40	88.9	216	10	Q9AXY2	Q9AXY2 brassica na
9	40	88.9	222	10	Q9AXY4	Q9AXY4 brassica ol
10	39	86.7	148	16	Q9RDS5	Q9RDS5 streptomyces
11	39	86.7	187	16	Q87ZM1	Q87ZM1 pseudomonas
12	39	86.7	293	16	Q91065	Q91065 pseudomonas
13	39	86.7	308	6	Q8WNV6	Q8WNV6 capra hircu
14	39	86.7	370	10	Q9M9G2	Q9M9G2 arabidopsis
15	39	86.7	407	10	Q8RXQ4	Q8RXQ4 arabidopsis
16	39	86.7	469	16	Q7UJM9	Q7UJM9 rhodospirillum

17	39	86.7	1404	16	Q824G8	Q824G8 chlamydomoni
18	38	84.4	70	4	Q8WXT4	Q8WXT4 homo sapien
19	38	84.4	178	6	Q95K65	Q95K65 macaca fasc
20	38	84.4	537	4	Q8WXT6	Q8WXT6 homo sapien
21	38	84.4	553	4	Q8WXT5	Q8WXT5 homo sapien
22	38	84.4	838	4	Q8WXT7	Q8WXT7 homo sapien
23	37	82.2	465	2	Q8GAG0	Q8GAG0 arthrobaacte
24	37	82.2	591	4	Q13041	Q13041 homo sapien
25	36	80.0	99	2	Q8GEF9	Q8GEF9 erwinia amy
26	36	80.0	115	10	Q8LI79	Q8LI79 oryza sativ
27	36	80.0	177	10	Q7XKD7	Q7XKD7 oryza sativ
28	36	80.0	185	16	Q9S581	Q9S581 pseudomonas
29	36	80.0	200	16	Q88IE8	Q88IE8 pseudomonas
30	36	80.0	223	16	Q98BA4	Q98BA4 rhizobium l
31	36	80.0	230	16	Q53604	Q53604 mycobacteri
32	36	80.0	230	16	Q7UZY3	Q7UZY3 mycobacteri
33	36	80.0	262	16	Q9HYG3	Q9HYG3 pseudomonas
34	36	80.0	264	16	Q8XZQ5	Q8XZQ5 ralatonia s
35	36	80.0	265	2	Q8KZQ7	Q8KZQ7 pseudomonas
36	36	80.0	265	2	Q85765	Q85765 pseudomonas
37	36	80.0	265	16	Q82B03	Q82B03 versinia pe
38	36	80.0	265	16	Q88R94	Q88R94 pseudomonas
39	36	80.0	286	16	Q8U649	Q8U649 agrobacteri
40	36	80.0	384	16	Q9RS74	Q9RS74 deinococcus
41	36	80.0	472	16	Q8YMN8	Q8YMN8 anabaena sp
42	36	80.0	475	16	Q8PFC2	Q8PFC2 xanthomonas
43	36	80.0	497	16	Q88HM3	Q88HM3 pseudomonas
44	36	80.0	699	16	Q82QU9	Q82QU9 streptomyce
45	36	80.0	754	16	Q8ZQC3	Q8ZQC3 salmonella
46	36	80.0	754	16	Q82B02	Q82B02 salmonella
47	36	80.0	774	11	Q8C771	Q8C771 mus musculu
48	36	80.0	1169	5	Q9VFX7	Q9VFX7 drosophila
49	36	80.0	1175	5	Q8WT43	Q8WT43 caenorhabdi
50	36	80.0	1178	5	Q8WT44	Q8WT44 caenorhabdi
51	36	80.0	1549	10	Q84K01	Q84K01 oryza sativ
52	36	80.0	1708	10	Q7XWZ9	Q7XWZ9 oryza sativ
53	35	77.8	108	16	Q7WLM9	Q7WLM9 bordetella
54	35	77.8	111	16	Q8ZPL2	Q8ZPL2 salmonella
55	35	77.8	111	16	Q8Z6S2	Q8Z6S2 salmonella
56	35	77.8	135	11	Q9CPW3	Q9CPW3 mus musculu
57	35	77.8	147	11	Q61427	Q61427 mus musculu
58	35	77.8	177	16	Q89X33	Q89X33 bradyrhizob
59	35	77.8	190	4	Q96N05	Q96N05 homo sapien
60	35	77.8	297	16	Q8EWI8	Q8EWI8 mycoplasma
61	35	77.8	302	2	Q8GJ78	Q8GJ78 mycobacteri
62	35	77.8	458	16	Q9PHQ5	Q9PHQ5 campylobact
63	35	77.8	494	9	Q7Y532	Q7Y532 bacterioph
64	35	77.8	494	17	Q974C6	Q974C6 sulfolobus
65	35	77.8	508	17	Q97WR4	Q97WR4 sulfolobus
66	35	77.8	522	5	Q9BHW0	Q9BHW0 leishmania
67	35	77.8	565	17	Q8PU00	Q8PU00 methanosarc
68	35	77.8	584	3	Q876A5	Q876A5 saccharomyc
69	35	77.8	621	5	Q95BS1	Q95BS1 leishmania
70	35	77.8	701	16	Q7VQA4	Q7VQA4 prochloroco
71	35	77.8	1004	16	Q8EKJ1	Q8EKJ1 shewanella
72	35	77.8	1092	10	Q84KB4	Q84KB4 cucumis mel
73	35	77.8	2898	3	Q872P1	Q872P1 neurospora
74	34	75.6	179	2	Q8G8Q0	Q8G8Q0 pseudomonas
75	34	75.6	179	16	Q91177	Q91177 pseudomonas
76	34	75.6	206	2	Q8GC77	Q8GC77 erwinia chr
77	34	75.6	207	16	Q7VDI5	Q7VDI5 prochloroco
78	34	75.6	217	16	Q8FC64	Q8FC64 escherichia
79	34	75.6	218	5	Q23570	Q23570 caenorhabdi
80	34	75.6	255	17	Q28806	Q28806 archaeoglob
81	34	75.6	283	2	Q8VW88	Q8VW88 terrabacter
82	34	75.6	301	2	Q52770	Q52770 pseudomonas
83	34	75.6	301	16	Q911B1	Q911B1 pseudomonas
84	34	75.6	301	16	Q882C1	Q882C1 pseudomonas
85	34	75.6	305	2	Q9S4I5	Q9S4I5 mycobacteri
86	34	75.6	305	2	Q9R2Z0	Q9R2Z0 mycobacteri
87	34	75.6	305	2	Q88998	Q88998 mycobacteri
88	34	75.6	306	2	Q933U0	Q933U0 mycobacteri
89	34	75.6	316	16	Q8X4U4	Q8X4U4 escherichia

90	34	75.6	316	16	Q8EGD3	Q8fgd3 escherichia	163	33	73.3	296	5	Q9W2M4	Q9w2m4 drosophila
91	34	75.6	322	5	Q8S2E2	Q8s2e2 drosophila	164	33	73.3	308	5	Q8MR08	Q8mr08 drosophila
92	34	75.6	329	16	Q8YZ09	Q8yz09 anabaena sp	165	33	73.3	313	16	Q82Xi9	Q82xi9 nitrosomona
93	34	75.6	339	16	Q8XYU6	Q8xyu6 ralstonia s	166	33	73.3	314	16	Q82Li20	Q82li20 streptomyces
94	34	75.6	366	16	Q8FC71	Q8fc71 streptomyces	167	33	73.3	316	5	Q62517	Q62517 caenorhabditis
95	34	75.6	373	11	Q8CA29	Q8ca29 mus musculus	168	33	73.3	320	2	Q93JR1	Q93jr1 rhodococcus
96	34	75.6	376	4	Q43898	Q43898 homo sapien	169	33	73.3	324	2	Q83Z76	Q83z76 citrobacter
97	34	75.6	384	16	Q7WJE4	Q7wje4 bordetella	170	33	73.3	329	16	Q83R78	Q83r78 bradyrhizob
98	34	75.6	384	16	Q7WAA1	Q7waa1 bordetella	171	33	73.3	345	2	Q8LOV9	Q8lov9 escherichia
99	34	75.6	384	16	Q7VYH3	Q7vyh3 bordetella	172	33	73.3	351	10	Q7X9S1	Q7x9s1 gossypium b
100	34	75.6	391	16	Q8CW5	Q8cw5 xylella fas	173	33	73.3	354	11	Q9CS82	Q9cs82 mus musculus
101	34	75.6	394	10	Q9AW80	Q9aw80 guillardia	174	33	73.3	361	10	Q9AUL3	Q9aul3 oryza sativ
102	34	75.6	396	16	Q87P2	Q87p2 xylella fas	175	33	73.3	364	10	Q8L936	Q8l936 arabidopsis
103	34	75.6	424	10	Q48663	Q48663 chlamydomon	176	33	73.3	365	10	Q9SMP9	Q9smp9 arabidopsis
104	34	75.6	425	2	Q83X73	Q83x73 streptomyces	177	33	73.3	376	2	Q31110	Q31110 pseudomonas
105	34	75.6	433	10	Q9M7J5	Q9m7j5 lophopyrum	178	33	73.3	377	16	Q8G242	Q8g242 bruceella su
106	34	75.6	454	10	Q9ZRW0	Q9zrw0 cicer ariet	179	33	73.3	381	2	Q93TQ4	Q93tq4 agrobacteri
107	34	75.6	471	16	Q87Z93	Q87z93 pseudomonas	180	33	73.3	386	16	Q8UH89	Q8uh89 agrobacteri
108	34	75.6	497	3	Q94514	Q94514 schizosacch	181	33	73.3	396	16	Q8ZMJ7	Q8zmj7 salmonella
109	34	75.6	499	10	Q9LTV0	Q9ltv0 arabidopsis	182	33	73.3	396	16	Q8X5A2	Q8x5a2 escherichia
110	34	75.6	501	16	Q82D02	Q82d02 streptomyces	183	33	73.3	396	16	Q8FFS3	Q8ffs3 escherichia
111	34	75.6	504	16	Q9KZR2	Q9kzr2 streptomyces	184	33	73.3	396	16	Q8Z581	Q8z581 salmonella
112	34	75.6	506	10	Q8GSQ6	Q8gsq6 lithospermu	185	33	73.3	396	16	Q83KE0	Q83ke0 shigella fl
113	34	75.6	510	16	Q8XQC5	Q8xqc5 ralstonia s	186	33	73.3	396	16	Q7UIG3	Q7uig3 rhodopirell
114	34	75.6	522	10	Q9SGT7	Q9sgt7 arabidopsis	187	33	73.3	403	16	Q8YFS6	Q8yfs6 bruceella me
115	34	75.6	525	16	Q8XQC8	Q8xqc8 ralstonia s	188	33	73.3	406	16	Q7UIX0	Q7uix0 rhodopirell
116	34	75.6	527	5	Q3VGF1	Q3vgt1 drosophila	189	33	73.3	407	16	Q9CG50	Q9cg50 lactococcus
117	34	75.6	545	16	Q87T14	Q87t14 vibrio para	190	33	73.3	421	10	Q9SB04	Q9sbu4 chlamydomon
118	34	75.6	549	10	Q8H2U5	Q8h2u5 oryza sativ	191	33	73.3	423	10	Q8H0T7	Q8hot7 arabidopsis
119	34	75.6	558	16	P73902	P73902 synchocyst	192	33	73.3	441	16	Q91IQ7	Q91iq7 pseudomonas
120	34	75.6	594	5	Q811I3	Q811i3 plasmodium	193	33	73.3	469	5	Q81J7	Q81j7 plasmodium
121	34	75.6	648	5	Q9ND58	Q9nds8 bombyx mori	194	33	73.3	470	16	Q8EXK0	Q8exk0 leptospira
122	34	75.6	668	16	Q81XZ2	Q81xz2 bacillus an	195	33	73.3	473	5	Q9BHU7	Q9bhv7 leishmania
123	34	75.6	782	16	Q81ZY4	Q81zy4 streptomyces	196	33	73.3	473	11	Q70396	Q70396 mus musculus
124	34	75.6	869	16	Q9CN28	Q9cn28 pasteurella	197	33	73.3	474	16	Q8F2K0	Q8fxk0 leptospira
125	34	75.6	910	17	O28621	O28621 archaeoglob	198	33	73.3	476	11	Q91XJ9	Q91xj9 mus musculus
126	34	75.6	2517	4	Q8N3W9	Q8n3w9 homo sapien	199	33	73.3	480	16	P74370	P74370 synchocyst
127	34	75.6	2517	4	Q81WT3	Q81wt3 homo sapien	200	33	73.3	483	16	Q82IT1	Q82it1 streptomyces
128	33	73.3	64	16	Q8RGN7	Q8rgn7 thermomanaer	201	33	73.3	484	16	Q88S64	Q88s64 lactobacill
129	33	73.3	70	11	Q7TSD7	Q7tsd7 rattus norv	202	33	73.3	485	10	O80401	O80401 oryza sativ
130	33	73.3	71	12	Q7T9H4	Q7t9h4 kaposi's sa	203	33	73.3	486	16	Q8KIC9	Q8kic9 fusobacteri
131	33	73.3	71	12	Q7T9H3	Q7t9h3 kaposi's sa	204	33	73.3	487	5	O45012	O45012 caenorhabdi
132	33	73.3	71	12	Q7T9H2	Q7t9h2 kaposi's sa	205	33	73.3	487	16	Q8ESK1	Q8esk1 leptospira
133	33	73.3	71	12	Q7T9H0	Q7t9h0 kaposi's sa	206	33	73.3	487	16	Q8ESK1	Q8esk1 oceanobacil
134	33	73.3	96	2	Q83ZH2	Q83zh2 streptococ	207	33	73.3	487	16	Q82ID4	Q82id4 streptomyces
135	33	73.3	106	2	Q57266	Q57266 bacillus an	208	33	73.3	489	13	Q7ZXX3	Q7zxx3 xenopus lae
136	33	73.3	106	16	Q9PE50	Q9pe50 xylella fas	209	33	73.3	494	16	Q8F3Z3	Q8fz3 leptospira
137	33	73.3	106	17	Q8UIG4	Q8uig4 pyrococcus	210	33	73.3	500	11	Q80Z27	Q80z27 mus musculus
138	33	73.3	110	2	O6062	O6062 marincoccus	211	33	73.3	502	16	Q8NTQ1	Q8ntq1 corynebacte
139	33	73.3	113	16	Q8KFP1	Q8kfp1 chlorobium	212	33	73.3	503	1	Q9C4N4	Q9c4n4 methanobrev
140	33	73.3	122	10	Q8SAC3	Q8sac3 oryza sativ	213	33	73.3	504	16	Q8CPD0	Q8cpd0 staphylococ
141	33	73.3	128	10	Q7XGL3	Q7xgl3 oryza sativ	214	33	73.3	505	2	Q9KWL9	Q9kwl9 staphylococ
142	33	73.3	128	16	Q88U91	Q88u91 pseudomonas	215	33	73.3	506	16	Q9EWJ4	Q9ewj4 streptomyces
143	33	73.3	130	9	Q8SB15	Q8sb15 streptococ	216	33	73.3	508	3	Q9P7S7	Q9p7s7 schizosacch
144	33	73.3	134	16	Q82X14	Q82x14 nitrosomona	217	33	73.3	510	5	Q9U5W4	Q9u5w4 drosophila
145	33	73.3	136	16	Q97Q08	Q97qg8 streptococ	218	33	73.3	511	5	Q9VM69	Q9vm69 drosophila
146	33	73.3	149	17	Q9H5F8	Q9h5f8 halobacteri	219	33	73.3	519	13	Q803P5	Q803p5 brachydanio
147	33	73.3	163	16	Q8DM71	Q8dm71 synchococ	220	33	73.3	533	10	Q9PPT6	Q9pt6 arabidopsis
148	33	73.3	175	16	Q8EIA5	Q8eias shewanella	221	33	73.3	533	10	Q9MAB3	Q9mab3 arabidopsis
149	33	73.3	212	16	Q8NP98	Q8np98 corynebacte	222	33	73.3	536	16	Q8CBY7	Q8cbz8 corynebacte
150	33	73.3	235	5	Q86AM7	Q86am7 dictyosteli	223	33	73.3	536	11	Q8CBY7	Q8cbz8 corynebacte
151	33	73.3	237	5	Q8T623	Q8t623 dictyosteli	224	33	73.3	550	10	O65335	O65335 pisum sativ
152	33	73.3	241	17	O59123	O59123 pyrococcus	225	33	73.3	555	10	Q9AV96	Q9av96 nicotiana t
153	33	73.3	249	5	Q9VA97	Q9va97 drosophila	226	33	73.3	560	10	O65334	O65334 pisum sativ
154	33	73.3	249	16	Q892F2	Q892f2 clostridium	227	33	73.3	563	16	Q87A6	Q87a6 vibrio vuln
155	33	73.3	253	2	Q9XDB7	Q9xdb7 methylococ	228	33	73.3	565	16	Q9KWL1	Q9kwl1 vibrio chol
156	33	73.3	254	5	Q8GP39	Q8gp39 drosophila	229	33	73.3	568	16	Q87HU2	Q87hu2 vibrio para
157	33	73.3	260	2	P97047	P97047 methylococ	230	33	73.3	597	3	Q8X066	Q8x066 neurospora
158	33	73.3	275	16	P74583	P74583 synchocyst	231	33	73.3	634	11	Q9D687	Q9d687 mus musculus
159	33	73.3	285	12	Q9WHC3	Q9whc3 kaposi's sa	232	33	73.3	722	11	Q7TQ67	Q7tq67 mus musculus
160	33	73.3	289	2	O05111	O05111 methylococ	233	33	73.3	722	11	Q7TN58	Q7tn58 mus musculus
161	33	73.3	289	12	Q9WHB7	Q9whb7 kaposi's sa	234	33	73.3	779	16	Q8YX39	Q8yx39 anabaena sp
162	33	73.3	290	16	Q82ZC6	Q82zc6 enterococ	235	33	73.3	779	16	Q8YX39	Q8yx39 anabaena sp

236	33	73.3	785	16	Q9CN05	Q9cm05 pasteurella	309	32	71.1	229	12	Q9WHB1	Q9whb1 kaposi's sa
237	33	73.3	809	16	Q9CNL9	Q9cnl9 pasteurella	310	32	71.1	233	15	Q8JBO5	Q8je05 human immun
238	33	73.3	818	16	Q8XTM1	Q8xtm1 ralstonia s	311	32	71.1	234	13	Q9W6U6	Q9w6u6 fugu rubrip
239	33	73.3	819	11	Q8BL40	Q8bl40 mus musculus	312	32	71.1	235	12	Q9IC40	Q9ic40 kaposi's sa
240	33	73.3	820	11	Q7TN22	Q7tn22 mus musculus	313	32	71.1	241	16	Q8DJZ4	Q8djz4 synchococce
241	33	73.3	862	16	Q8E9C6	Q8e9c6 shewanella	314	32	71.1	258	16	Q8NLA5	Q8nla5 corynebacte
242	33	73.3	1015	5	Q8MRT3	Q8mrt3 drosophila	315	32	71.1	260	16	Q83250	Q83250 treponema p
243	33	73.3	1015	16	Q8ZPE5	Q8zpe5 salmonella	316	32	71.1	261	16	Q88F09	Q88f09 pseudomonas
244	33	73.3	1015	16	Q8XAS2	Q8xas2 escherichia	317	32	71.1	264	16	Q83BG2	Q83bg2 coxiella bu
245	33	73.3	1015	16	Q8FHH8	Q8fhh8 escherichia	318	32	71.1	266	12	Q9IC88	Q9ic88 kaposi's sa
246	33	73.3	1015	16	Q83R95	Q83r95 shigella fl	319	32	71.1	267	12	Q9WH9A	Q9wh9a kaposi's sa
247	33	73.3	1016	16	Q8ZKS8	Q8zks8 salmonella	320	32	71.1	270	5	Q8ITC0	Q8itc0 aequipecten
248	33	73.3	1016	16	Q8XBB7	Q8xbb7 escherichia	321	32	71.1	270	10	Q84XU3	Q84xu3 elaeis guin
249	33	73.3	1016	16	Q8FBE7	Q8fbe7 escherichia	322	32	71.1	272	12	Q9IC30	Q9ic30 kaposi's sa
250	33	73.3	1016	16	Q8Z2U6	Q8z2u6 salmonella	323	32	71.1	272	12	Q9WH9A	Q9wh9a kaposi's sa
251	33	73.3	1016	16	Q83FE6	Q83fe6 shigella fl	324	32	71.1	272	12	Q9IC28	Q9ic28 kaposi's sa
252	33	73.3	1024	16	Q8DMW3	Q8dmw3 synchococce	325	32	71.1	273	12	Q995B7	Q995b7 kaposi's sa
253	33	73.3	1034	16	Q931E0	Q931e0 rhizobium m	326	32	71.1	273	16	Q8K6H9	Q8k6h9 streptococ
254	33	73.3	1164	10	Q81517	Q81517 arabidopsis	327	32	71.1	275	12	Q995A4	Q995a4 kaposi's sa
255	33	73.3	1284	2	Q8RQU9	Q8rxu9 bacillus gl	328	32	71.1	275	12	Q9IC26	Q9ic26 kaposi's sa
256	33	73.3	1368	16	Q8XSS7	Q8xsj7 ralstonia s	329	32	71.1	276	12	Q9IC27	Q9ic27 kaposi's sa
257	33	73.3	1641	8	Q9TAL3	Q9tal3 catereeria r	330	32	71.1	276	12	Q9IC31	Q9ic31 kaposi's sa
258	33	73.3	1956	4	Q9Y5Y9	Q9y5y9 homo sapien	331	32	71.1	276	12	Q9WNS7	Q9wns7 kaposi's sa
259	33	73.3	1962	6	Q46669	Q46669 canis famil	332	32	71.1	276	12	Q9IC25	Q9ic25 kaposi's sa
260	33	73.3	2060	5	Q7YUN7	Q7yun7 trypanosoma	333	32	71.1	276	17	Q8PUY8	Q8pu8 methanosaic
261	33	73.3	2060	5	Q7YUN5	Q7yun5 trypanosoma	334	32	71.1	279	12	Q995B6	Q995b6 kaposi's sa
262	33	73.3	2812	3	Q74630	Q74630 schizosacch	335	32	71.1	279	12	Q995A3	Q995a3 kaposi's sa
263	33	73.3	2846	2	Q8RL76	Q8rl76 pseudomonas	336	32	71.1	280	12	Q9WH98	Q9wh98 kaposi's sa
264	32.5	72.2	569	2	Q9LC92	Q9lc92 bacillus ha	337	32	71.1	280	12	Q9WH98	Q9wh98 kaposi's sa
265	32.5	72.2	572	16	Q9KCF7	Q9kcf7s bacillus ha	338	32	71.1	280	12	Q9WHB2	Q9whb2 kaposi's sa
266	32	71.1	71	12	Q9IC04	Q9ic04 kaposi's sa	339	32	71.1	280	16	Q9S284	Q9s284 streptomyc
267	32	71.1	71	12	Q9IC20	Q9ic20 kaposi's sa	340	32	71.1	281	12	Q91GV0	Q91gv0 kaposi's sa
268	32	71.1	71	12	Q9IBZ4	Q9ibz4 kaposi's sa	341	32	71.1	285	12	Q9WHH9	Q9whh9 kaposi's sa
269	32	71.1	71	12	Q9IC23	Q9ic23 kaposi's sa	342	32	71.1	288	16	Q8AOE4	Q8aoe4 bacteroides
270	32	71.1	71	12	Q91112	Q91112 kaposi's sa	343	32	71.1	289	12	Q9WH95	Q9wh95 kaposi's sa
271	32	71.1	71	12	Q9IBZ6	Q9ibz6 kaposi's sa	344	32	71.1	289	12	Q9DSB8	Q9dsb8 kaposi's sa
272	32	71.1	71	12	Q9IBZ7	Q9ibz7 kaposi's sa	345	32	71.1	289	12	Q9DSF5	Q9dsf5 kaposi's sa
273	32	71.1	71	12	Q9IBZ8	Q9ibz8 kaposi's sa	346	32	71.1	289	12	Q9IC84	Q9ic84 kaposi's sa
274	32	71.1	71	12	Q9IC03	Q9ic03 kaposi's sa	347	32	71.1	289	12	Q9DSC5	Q9dsc5 kaposi's sa
275	32	71.1	71	12	Q80P27	Q80p27 kaposi's sa	348	32	71.1	289	12	Q9WHC1	Q9whc1 kaposi's sa
276	32	71.1	71	12	Q80P26	Q80p26 kaposi's sa	349	32	71.1	289	12	Q9DSR4	Q9dsr4 kaposi's sa
277	32	71.1	71	12	Q80P25	Q80p25 kaposi's sa	350	32	71.1	289	12	Q9DSR0	Q9dsr0 kaposi's sa
278	32	71.1	90	16	Q7UPJ0	Q7upj0 rhodopirell	351	32	71.1	289	12	Q9WHC0	Q9whc0 kaposi's sa
279	32	71.1	106	2	Q7WJX7	Q7wxj7 alcaligenes	352	32	71.1	289	12	Q9WHB9	Q9whb9 kaposi's sa
280	32	71.1	110	16	Q9JZV7	Q9jzv7 neisseria m	353	32	71.1	289	12	P88970	P88970 kaposi's sa
281	32	71.1	110	16	Q9JUX0	Q9jux0 neisseria m	354	32	71.1	289	12	P88970	P88970 kaposi's sa
282	32	71.1	114	12	Q995C0	Q995c0 kaposi's sa	355	32	71.1	289	12	Q9IC87	Q9ic87 kaposi's sa
283	32	71.1	115	12	Q995A1	Q995a1 kaposi's sa	356	32	71.1	289	12	Q9WHH8	Q9whh8 kaposi's sa
284	32	71.1	116	2	Q8GNK0	Q8gnk0 mycobacteri	357	32	71.1	289	12	Q9DSF1	Q9dsf1 kaposi's sa
285	32	71.1	118	12	Q995C2	Q995c2 kaposi's sa	358	32	71.1	289	12	Q9WHC7	Q9whc7 kaposi's sa
286	32	71.1	121	12	Q995C3	Q995c3 kaposi's sa	359	32	71.1	289	12	Q9IC89	Q9ic89 kaposi's sa
287	32	71.1	121	12	Q995B8	Q995b8 kaposi's sa	360	32	71.1	289	12	Q80IF6	Q80if6 kaposi's sa
288	32	71.1	123	4	Q9NKN1	Q9nkn1 homo sapien	361	32	71.1	289	12	Q80IF4	Q80if4 kaposi's sa
289	32	71.1	126	2	Q8VM97	Q8vm97 alcaligenes	362	32	71.1	289	12	Q80IF3	Q80if3 kaposi's sa
290	32	71.1	129	12	Q7TFV7	Q7tfv7 thesus cyto	363	32	71.1	289	12	Q80IF2	Q80if2 kaposi's sa
291	32	71.1	134	13	Q90W69	Q90w69 oncorhynch	364	32	71.1	289	12	Q80IF1	Q80if1 kaposi's sa
292	32	71.1	135	4	Q9H5Z4	Q9h5z4 homo sapien	365	32	71.1	289	12	Q80IF0	Q80if0 kaposi's sa
293	32	71.1	141	4	Q9NU66	Q9nu66 homo sapien	366	32	71.1	289	12	Q80IE7	Q80ie7 kaposi's sa
294	32	71.1	141	13	Q90XHL	Q90xhl oncorhynch	367	32	71.1	289	12	Q80IE5	Q80ie5 kaposi's sa
295	32	71.1	143	6	Q9GKV9	Q9gkv9 macaca fasc	368	32	71.1	289	12	Q80IE2	Q80ie2 kaposi's sa
296	32	71.1	149	16	Q8A9C8	Q8a9c8 bacteroides	369	32	71.1	289	12	Q80IE1	Q80ie1 kaposi's sa
297	32	71.1	154	16	Q8ZGG2	Q8zgg2 versinia pe	370	32	71.1	289	12	Q80ID7	Q80id7 kaposi's sa
298	32	71.1	158	2	Q9RIK3	Q9rik3 streptococ	371	32	71.1	289	12	Q80ID5	Q80id5 kaposi's sa
299	32	71.1	159	16	Q92NC6	Q92nc6 rhizobium m	372	32	71.1	289	12	Q80ID6	Q80id6 kaposi's sa
300	32	71.1	171	16	Q7UYX1	Q7uyx1 rhodopirell	373	32	71.1	289	12	Q80ID3	Q80id3 kaposi's sa
301	32	71.1	194	16	Q82W50	Q82w50 nitrosonoma	374	32	71.1	289	12	Q80ID5	Q80id5 kaposi's sa
302	32	71.1	198	2	Q8GE39	Q8ge39 heliobacill	375	32	71.1	289	12	Q80SH3	Q80sh3 pseudomonas
303	32	71.1	206	5	Q8MUE7	Q8mue7 chlamys far	376	32	71.1	289	16	Q9HYA0	Q9hya0 pseudomonas
304	32	71.1	209	16	Q8FTW7	Q8ftw7 corynebacte	377	32	71.1	289	17	Q8TRL3	Q8trl3 methanosaic
305	32	71.1	210	16	Q9K8S1	Q9k8s1 bacillus ha	378	32	71.1	290	12	Q9DSF2	Q9dsf2 kaposi's sa
306	32	71.1	225	16	Q88IV8	Q88iv8 pseudomonas	379	32	71.1	290	12	Q9IC90	Q9ic90 kaposi's sa
307	32	71.1	227	12	Q9WH9A	Q9wh9a kaposi's sa	380	32	71.1	290	12	Q80IE6	Q80ie6 kaposi's sa
308	32	71.1	228	12	Q9W9W9	Q9w9w9 kaposi's sa	381	32	71.1	290	12	Q80IE4	Q80ie4 kaposi's sa

382	32	71.1	295	3	Q8J244	Q8J244 phaeosphaera	455	32	71.1	474	16	Q8F3V9	Q8F3v9 leptospira
383	32	71.1	300	3	Q8U240	Q8U240 mycosphaera	455	32	71.1	477	16	Q8YX0	Q8YPY0 anabaena sp
384	32	71.1	300	10	Q8XIF7	Q8XIF7 arabidopsis	457	32	71.1	477	16	Q8XZM5	Q8XZM5 ralsstonia s
385	32	71.1	300	16	Q8UW70	Q8UW70 lactobacill	458	32	71.1	478	16	Q8F3B8	Q8F3B8 leptospira
386	32	71.1	302	12	Q8OLM2	Q8OLM2 adoxophyes	459	32	71.1	478	16	Q8XDB8	Q8XDB8 bradyrhizob
387	32	71.1	307	10	Q8ASS3	Q8ASS3 arabidopsis	460	32	71.1	478	16	Q8VKA9	Q8VKA9 helicobacte
388	32	71.1	309	5	Q8GGB8	Q8GGB8 caenorhabdi	461	32	71.1	479	16	Q8VGA9	Q8VGA9 vibrio para
389	32	71.1	325	11	Q8CSM5	Q8CSM5 mus musculu	462	32	71.1	480	16	Q8ZGS4	Q8ZGS4 yersinia pe
390	32	71.1	326	10	Q8LQ43	Q8LQ43 arabidopsis	463	32	71.1	483	2	Q8AQQ9	Q8AQQ9 actinobacil
391	32	71.1	327	2	Q8ABF0	Q8ABF0 xanthomonas	464	32	71.1	484	2	Q8RG14	Q8RG14 rhizobium s
392	32	71.1	336	16	Q8V9L9	Q8V9L9 prochloroco	465	32	71.1	484	2	P77939	P77939 pasteurella
393	32	71.1	338	10	Q8H0W9	Q8H0W9 arabidopsis	466	32	71.1	484	16	Q8CPK5	Q8CPK5 leptospira
394	32	71.1	338	10	Q8LX94	Q8LX94 arabidopsis	467	32	71.1	484	16	Q8F6U7	Q8F6U7 leptospira
395	32	71.1	342	16	Q87713	Q87713 aquifex seo	468	32	71.1	485	16	Q82TK1	Q82TK1 nitrosomona
396	32	71.1	363	10	Q8AFV3	Q8AFV3 chlamydomon	469	32	71.1	485	16	Q81UI2	Q81UI2 bacillus an
397	32	71.1	367	17	Q8ZWC3	Q8ZWC3 pyrobaculum	470	32	71.1	486	16	Q8EHY4	Q8EHY4 shewanella
398	32	71.1	368	5	Q8BHE1	Q8BHE1 leishmania	471	32	71.1	495	16	Q83577	Q83577 treponema p
399	32	71.1	375	13	Q8ZYR6	Q8ZYR6 xenopus lae	472	32	71.1	495	16	Q8A5C2	Q8A5C2 bacteroides
400	32	71.1	375	13	Q8ZW95	Q8ZW95 brachydanio	473	32	71.1	496	16	Q81674	Q81674 pseudomonas
401	32	71.1	375	16	Q8WF99	Q8WF99 bordetella	474	32	71.1	499	16	Q8YMR6	Q8YMR6 anabaena sp
402	32	71.1	375	16	Q8W3W9	Q8W3W9 bordetella	475	32	71.1	499	16	Q8FVU0	Q8FVU0 brucella su
403	32	71.1	376	16	Q8VUT6	Q8VUT6 bordetella	476	32	71.1	499	16	Q82VM1	Q82VM1 nitrosomona
404	32	71.1	376	16	Q8Z6X9	Q8Z6X9 listeria in	477	32	71.1	503	16	Q8CBV2	Q8CBV2 mycobacteri
405	32	71.1	376	16	Q8Y3R5	Q8Y3R5 listeria mo	478	32	71.1	503	16	Q83944	Q83944 mycobacteri
406	32	71.1	383	16	Q8ZGR3	Q8ZGR3 streptomyce	479	32	71.1	503	16	Q87Z19	Q87Z19 mycobacteri
407	32	71.1	385	13	Q8OYW7	Q8OYW7 ictalurus p	480	32	71.1	504	16	Q8JRF5	Q8JRF5 neisseria m
408	32	71.1	388	16	Q81T98	Q81T98 bacillus an	481	32	71.1	505	2	Q814S1	Q814S1 staphylococ
409	32	71.1	391	16	Q81G40	Q81G40 bacillus ce	482	32	71.1	505	16	Q8NWX5	Q8NWX5 staphylococ
410	32	71.1	394	16	Q8ZD48	Q8ZD48 listeria in	483	32	71.1	506	2	Q83WC7	Q83WC7 acinetobact
411	32	71.1	394	16	Q8S390	Q8S390 listeria mo	484	32	71.1	507	10	Q84VU3	Q84VU3 solenostemo
412	32	71.1	396	13	Q8ZY88	Q8ZY88 xenopus lae	485	32	71.1	507	16	Q89UE2	Q89UE2 staphylococ
413	32	71.1	400	16	Q83ON1	Q83ON1 enterococcu	486	32	71.1	509	10	Q8L5H7	Q8L5H7 ocimum basi
414	32	71.1	401	13	Q8SY96	Q8SY96 xenopus lae	487	32	71.1	527	16	Q82526	Q82526 helicobacte
415	32	71.1	403	10	Q8VYD7	Q8VYD7 arabidopsis	488	32	71.1	531	8	Q8GCB8	Q8GCB8 zea mays (m
416	32	71.1	404	2	Q8XZM5	Q8XZM5 staphylococ	489	32	71.1	548	16	Q82AK2	Q82AK2 streptomyce
417	32	71.1	404	10	Q8XLI5	Q8XLI5 oryza sativ	490	32	71.1	548	5	Q80963	Q80963 leishmania
418	32	71.1	404	16	Q8CFW1	Q8CFW1 staphylococ	491	32	71.1	549	5	Q8FK65	Q8FK65 arabidopsis
419	32	71.1	404	16	Q8AVM7	Q8AVM7 lactobacill	492	32	71.1	570	16	Q8V6R1	Q8V6R1 prochloroco
420	32	71.1	404	16	Q83662	Q83662 staphylococ	493	32	71.1	576	4	Q86T67	Q86T67 homo sapien
421	32	71.1	406	10	Q8LA93	Q8LA93 arabidopsis	494	32	71.1	578	4	Q86451	Q86451 saccharomyc
422	32	71.1	407	10	Q8GHH3	Q8GHH3 arabidopsis	495	32	71.1	622	3	Q820D6	Q820D6 navel orang
423	32	71.1	413	16	Q87N83	Q87N83 streptococ	496	32	71.1	699	12	Q8Z0D6	Q8Z0D6 prochloroco
424	32	71.1	414	16	Q8DN12	Q8DN12 streptococ	497	32	71.1	704	16	Q8V5Y7	Q8V5Y7 prochloroco
425	32	71.1	415	16	Q8VH38	Q8VH38 helicobacte	498	32	71.1	721	10	Q8LTA0	Q8LTA0 arabidopsis
426	32	71.1	418	13	Q8IBG2	Q8IBG2 gallus gall	499	32	71.1	738	10	Q8RWS6	Q8RWS6 arabidopsis
427	32	71.1	418	16	Q89Z27	Q89Z27 streptococ	500	32	71.1	739	10	Q8LT99	Q8LT99 arabidopsis
428	32	71.1	418	16	Q8P0K0	Q8P0K0 streptococ		32	71.1	775	10	Q8LG72	Q8LG72 oryza sativ
429	32	71.1	418	16	Q8K745	Q8K745 streptococ							
430	32	71.1	419	13	Q8AXV3	Q8AXV3 fugu rubrip							
431	32	71.1	419	16	Q8UNA5	Q8UNA5 rhodopirell							
432	32	71.1	420	2	Q80067	Q80067 streptococ							
433	32	71.1	420	16	Q86577	Q86577 streptococ							
434	32	71.1	421	2	Q8VM66	Q8VM66 streptococ							
435	32	71.1	421	16	Q8E3C8	Q8E3C8 streptococ							
436	32	71.1	421	16	Q8DXQ9	Q8DXQ9 streptococ							
437	32	71.1	432	17	Q8U3M1	Q8U3M1 pyrococcus							
438	32	71.1	433	10	Q8MLW1	Q8MLW1 arabidopsis							
439	32	71.1	435	10	Q851W9	Q851W9 oryza sativ							
440	32	71.1	441	16	Q81M20	Q81M20 bacillus an							
441	32	71.1	441	16	Q81B22	Q81B22 bacillus ce							
442	32	71.1	457	5	Q860D1	Q860D1 drosophila							
443	32	71.1	465	16	Q8EXJ8	Q8EXJ8 leptospira							
444	32	71.1	466	5	Q17451	Q17451 culex pipie							
445	32	71.1	470	2	Q49133	Q49133 methylobact							
446	32	71.1	470	16	Q897X8	Q897X8 clostridium							
447	32	71.1	470	16	Q8VH83	Q8VH83 helicobacte							
448	32	71.1	470	16	Q8VH83	Q8VH83 helicobacte							
449	32	71.1	471	16	Q83MR9	Q83MR9 myxococcus							
450	32	71.1	471	16	Q888J1	Q888J1 pseudomonas							
451	32	71.1	471	16	Q81U73	Q81U73 bacillus an							
452	32	71.1	472	11	Q811T8	Q811T8 mus musculu							
453	32	71.1	473	16	Q871S6	Q871S6 clostridium							
454	32	71.1	473	16	Q8F4B7	Q8F4B7 leptospira							

ALIGNMENTS

RESULT 1

Q88QK9	PRELIMINARY;	PRT;	479 AA.
ID	Q88QK9	PRELIMINARY;	PRT;
AC	Q88QK9;	(TREMBlrel. 24, Created)	
DT	01-JUN-2003	(TREMBlrel. 24, Last sequence update)	
DT	01-JUN-2003	(TREMBlrel. 24, Last sequence update)	
DT	01-OCT-2003	(TREMBlrel. 25, Last annotation update)	
DE	Catalase.		
GN	KATA OR PP0481.		
OS	Pseudomonas putida (strain KT2440).		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;		
OC	Pseudomonadaceae; Pseudomonas.		
OX	NCBI_TaxID=160488;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22432060; PubMed=12534463;		
RA	Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,		
RA	Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,		
RA	Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,		
RA	Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,		
RA	Chris Lee P., Holtzapflee E., Scanlan D., Tran K., Moazzes A.,		
RA	Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,		

RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile *Pseudomonas putida* KT2440.";
 RL Environ. Microbiol. 4:799-808 (2002).
 DR EMBL; AB016775; AAM6111.1; -.
 DR TIGR; PP0481; -.
 DR GO; GO:0004096; F:catalase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006979; P:response to oxidative stress; IEA.
 DR InterPro; IPR002226; Catalase.
 DR Pfam; PF00199; catalase.1.
 DR PRINTS; PR00067; CATALASE.
 DR PROSITE; PS00437; CATALASE_1; 1.
 DR PROSITE; PS00438; CATALASE_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 479 AA; 53589 MW; F874B178BD3DA501 CRC64;

 Query Match 93.3%; Score 42; DB 16; Length 479;
 Best Local Similarity 83.3%; Pred. No. 52;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 Qy 1 VWRWHF 6
 Db 209 WVKWHF 214

 RESULT 2
 Q9FIN2
 ID Q9FIN2 PRELIMINARY; PRT; 256 AA.
 AC Q9FIN2
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-JUN-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE 6-phosphogluconolactonase-like protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosid II; Brassicales; Brassicaceae; Arabidopsi
 OX NCBI_TaxID=3702;
 RN [1]
 RC STRAIN=Columbia;
 RX MEDLINE=99156233; PubMed=10048488;
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:379-391 (1998).
 DR EMBL; AB016884; BAB1233.1; -.
 DR GO; GO:0017057; F:6-phosphogluconolactonase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006148; Gluc_gal_isom.
 DR InterPro; IPR005900; Phosphogluconac.
 DR Pfam; PF01182; Glucosamine_1so; 1.
 DR TIGRFA; TIGR01198; pgl; 1.
 SQ SEQUENCE 256 AA; 28034 MW; 21E7046ECFDC72AF CRC64;

 Query Match 91.1%; Score 41; DB 10; Length 256;
 Best Local Similarity 83.3%; Pred. No. 40;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 Qy 1 VWRWHF 6
 Db 72 WARWHF 77

 RESULT 3
 Q8LD77
 ID Q8LD77 PRELIMINARY; PRT; 325 AA.

AC Q8LD77;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE 6-phosphogluconolactonase-like protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosid II; Brassicales; Brassicaceae; Arabidopsi
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0 (2002).
 RN [2]
 RC SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY086161; AAM63366.1; -.
 DR GO; GO:0017057; F:6-phosphogluconolactonase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006148; Gluc_gal_isom.
 DR InterPro; IPR005900; Phosphogluconac.
 DR Pfam; PF01182; Glucosamine_1so; 1.
 DR TIGRFA; TIGR01198; pgl; 1.
 SQ SEQUENCE 325 AA; 35592 MW; C51D8544114943D2 CRC64;

 Query Match 91.1%; Score 41; DB 10; Length 325;
 Best Local Similarity 83.3%; Pred. No. 51;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 Qy 1 VWRWHF 6
 Db 141 WARWHF 146

 RESULT 4
 Q84WW2
 ID Q84WW2 PRELIMINARY; PRT; 325 AA.
 AC Q84WW2
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative 6-phosphogluconolactonase.
 GN AT5G24400.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosid II; Brassicales; Brassicaceae; Arabidopsi
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,
 RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
 RA Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RT "Arabidopsis Open Reading Frame (ORF) Clones.";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT001923; AAN71922.1; -.
 DR GO; GO:0017057; F:6-phosphogluconolactonase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006148; Gluc_gal_isom.
 DR InterPro; IPR005900; Phosphogluconac.
 DR Pfam; PF01182; Glucosamine_1so; 1.
 DR TIGRFA; TIGR01198; pgl; 1.


```

SQ SEQUENCE 325 AA; 35644 MW; 0FBC6E95F9C073DC CRC64;
Query Match 91.1%; Score 41; DB 10; Length 325;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 141 WVRWHF 146

RESULT 5
Q9C600
ID Q9C600 PRELIMINARY; PRT; 201 AA.
AC Q9C600;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (Fragment).
GN FAD6
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsais.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Fourmann M., Froger N., Brunel D.;
RT "Amplified consensus gene markers: Tools designing for a genetic map
of Arabidopsis-known-function genes in Brassica.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229392; AAK00564.1; -.
GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase; 1.
DR ProDom; PD001081; FA_desat_fam; 1.
FT NON_TER 1
FT NON_TER 201
SQ SEQUENCE 201 AA; 23084 MW; 9261EEC0ED3771C0 CRC64;

Query Match 88.9%; Score 40; DB 10; Length 201;
Best Local Similarity 83.3%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 156 WVRWHF 161

RESULT 6
Q9AXY3
ID Q9AXY3 PRELIMINARY; PRT; 202 AA.
AC Q9AXY3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (Fragment).
GN FAD6-BR-1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. R500;
RA Fourmann M., Froger N., Brunel D.;
RT "Amplified consensus gene markers: Tools designing for a genetic map
of Arabidopsis-known-function genes in Brassica.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229390; AAK00662.1; -.
GO; GO:0016491; F:oxidoreductase activity; IEA.

```

```

DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase; 1.
DR ProDom; PD001081; FA_desat_fam; 1.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23163 MW; A96E403AAB1EC008 CRC64;

Query Match 88.9%; Score 40; DB 10; Length 202;
Best Local Similarity 83.3%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 157 WVRWHF 162

RESULT 7
Q9AXY5
ID Q9AXY5 PRELIMINARY; PRT; 213 AA.
AC Q9AXY5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (Fragment).
GN FAD6-BN-2.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Stellar;
RA Fourmann M., Froger N., Brunel D.;
RT "Amplified consensus gene markers: Tools designing for a genetic map
of Arabidopsis-known-function genes in Brassica.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229388; AAK00660.1; -.
GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase; 1.
DR ProDom; PD001081; FA_desat_fam; 2.
FT NON_TER 1
FT NON_TER 213
SQ SEQUENCE 213 AA; 24771 MW; D30EB4E98ADBD0C6 CRC64;

Query Match 88.9%; Score 40; DB 10; Length 213;
Best Local Similarity 83.3%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 154 WVRWHF 159

RESULT 8
Q9AXY2
ID Q9AXY2 PRELIMINARY; PRT; 216 AA.
AC Q9AXY2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (Fragment).
GN FAD6-BN-1.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Stellar;
RA Fourmann M., Froger N., Brunel D.;

```

"Amplified consensus gene markers: Tools designing for a genetic map of Arabidopsis known-function genes in Brassica.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF229391; AAK00663.1; -.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR InterPro; IPR005804; FA desat fam.
 DR Pfam; PF00487; FA desaturase; 1.
 DR ProDom; PD001081; FA desat_fam; 2.
 FT NON_TER 1 216
 FT NON_TER 216 216
 SQ SEQUENCE 216 AA; 25110 MW; EFPCEEDBBIB0E3E4 CRC64;

Query Match 88.9%; Score 40; DB 10; Length 216;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 |||||
 Db 157 WVNWHF 162

RESULT 9

Q9AXY4 PRELIMINARY; PRT; 222 AA.

ID Q9AXY4
 AC Q9AXY4
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Omega 6 reductase (Fragment).
 GN FAD6-BO-1.
 OS Brassica oleracea (Cauliflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3712;
 RN [1]

SEQUENCE FROM N.A.

RC STRAIN=cv. Rapid Cycling;
 RA Fourmann M., Froger N., Brunel D.;
 RA "Amplified consensus gene markers: Tools designing for a genetic map of Arabidopsis known-function genes in Brassica.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF229389; AAK00661.1; -.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR InterPro; IPR005804; FA desat fam.
 DR Pfam; PF00487; FA desaturase; 1.
 DR ProDom; PD001081; FA desat_fam; 2.
 FT NON_TER 1 222
 FT NON_TER 222 222
 SQ SEQUENCE 222 AA; 25761 MW; A2B19F83893BCD4E CRC64;

Query Match 88.9%; Score 40; DB 10; Length 222;
 Best Local Similarity 83.3%; Pred. No. 51;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 |||||
 Db 163 WVNWHF 168

RESULT 10

Q9RDS5 PRELIMINARY; PRT; 148 AA.

ID Q9RDS5
 AC Q9RDS5
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative membrane protein.
 GN SCO2242 OR SCIG2.04C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RA "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147 (2002).
 DR EMBL; AL939111; CAB66176.1; -.
 KW Complete proteome.
 SQ SEQUENCE 148 AA; 16191 MW; C7212463DF1B1D5C CRC64;

Query Match 86.7%; Score 39; DB 16; Length 148;
 Best Local Similarity 66.7%; Pred. No. 49;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 |||||
 Db 125 WVRWHF 130

RESULT 11

Q87ZM1 PRELIMINARY; PRT; 187 AA.

ID Q87ZM1
 AC Q87ZM1
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Baseplate assembly protein V.
 GN PSPT03403.
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Utterback T., Van Aken S., Feidblyum T., Gwin M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
 RA White O., Fraser C., Collmer A.;
 RA "Complete sequence of Pseudomonas syringae";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016868; AAO56881.1; -.
 DR TIGR; PSPT03403; -.
 DR InterPro; IPR006531; Phage_P2_V.
 DR Pfam; PF04717; phage_base_V; 1.
 KW Complete proteome.
 SQ SEQUENCE 187 AA; 19493 MW; DDEF0F4AF451C414 CRC64;

Query Match 86.7%; Score 39; DB 16; Length 187;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 5
 |||||
 Db 44 WVRWHF 48

RESULT 12

Q9I065 PRELIMINARY; PRT; 293 AA.

ID Q9I065
 AC Q9I065
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)

```

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PA2778.
GN
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004705; AG061661.1; -.
DR PIR; A83299; A83299.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR005074; peptidase_C39.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF03412; Peptidase_C39; 1.
DR Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 293 AA; 32016 MW; 06A410D6EDF60DE1 CRC64;

Query Match 86.7%; Score 39; DB 16; Length 293;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
DB 101 WPRWHF 106

RESULT 13
Q8WNV6 ID Q8WNV6 PRELIMINARY; PRT; 308 AA.
AC Q8WNV6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative pheromone receptor gVIR1.
GN gVIR1.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Shiba;
RX Wakabayashi Y., Mori Y., Ichikawa M., Yazaki K., Hagino-Yamagishi K.;
RT "gene of goat putative pheromone receptor is expressed in two distinct
RT olfactory organs.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB064662; BAB83523.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016503; F:pheromone receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:rhodopsin-like receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR004072; Vmron_receptorl.
DR Pfam; PF03402; VIR; 1.

DR PRINTS; PR01534; VOMERONASLIR.
DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 308 AA; 35370 MW; C86A6D3F41420192 CRC64;

Query Match 86.7%; Score 39; DB 6; Length 308;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 304 WVRWH 308

RESULT 14
Q9M9G2 ID Q9M9G2 PRELIMINARY; PRT; 370 AA.
AC Q9M9G2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F14023.22 protein.
GN F14023.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. columbia;
RA Liu S., Yu G., Lee J., Sakano H., Jhaveri A., Lenz C., Toriumi M.,
RA Chin C., Chiou J., Choi E., Gonzalez A., Hoang B., Koo T., Li J.,
RA Liu A., Pham P., Vaysberg M., Altati K., Buehler E., Chao Q., Conn L.,
RA Conway A.B., Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Nguyen M., Palm C., Shinn P., Tambunga G., Davis R., Ecker J.,
RA Federspiel N., Theologis A.;
RT "The sequence of BAC F14023 from Arabidopsis thaliana chromosome 1.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. columbia;
RA Theologis A.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012654; AAF43237.1; -.
DR PIR; A96741; A96741.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 3.
DR PROSITE; PS00082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
DX Repeat; WD repeat.
KW Repeat; WD repeat.
SQ SEQUENCE 370 AA; 39774 MW; CE58B01F850C3B13 CRC64;

Query Match 86.7%; Score 39; DB 10; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 151 WVRWH 155

RESULT 15
Q8RXQ4 ID Q8RXQ4 PRELIMINARY; PRT; 407 AA.
AC Q8RXQ4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN ATIG71840.

```

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA "Arabidopsis Full Length cDNA Clones.";
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Huan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shim P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY080732; AAL86002.1; -;
DR EMBL; BT001964; AAN71963.1; -;
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 8
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 8.
DR PROSITE; PS00678; WD REPEATS 1; 3.
DR PROSITE; PS00082; WD REPEATS 2; 8.
DR PROSITE; PS00294; WD REPEATS REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 407 AA; 43791 MW; 712D7CADD1790833 CRC64;

Query Match 86.7%; Score 39; DB 10; Length 407;
Best Local Similarity 100.0%; Pred.No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 160 WVRWH 164

RESULT 16
QYU09 PRELIMINARY; PRT; 469 AA.
ID QYU09
AC QYU09
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Hypothetical protein.
GN RB11174.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]_TaxID=117;
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294152; CAD77203.1; -;

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 469 AA; 51193 MW; 02B4065710CB2A04 CRC64;

Query Match 86.7%; Score 39; DB 16; Length 469;
Best Local Similarity 83.3%; Pred.No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 6
DB 333 WVRWH 338

RESULT 17
QYU09 PRELIMINARY; PRT; 1404 AA.
ID QYU09
AC QYU09
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE SecDF protein, putative.
GN CCA00178.
OS Chlamydomonas reinhardtii.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J., Holtzapple E., Khouri H., Federova N.B., Carty H.A.,
RA Umayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
RA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavovill P.M.,
RA Fraser C.M.;
RT "Genome sequence of Chlamydomonas reinhardtii (Chlamydia GPIC).
examining the role of niche-specific genes in the evolution of the
Chlamydiae.";
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016994; AAP04929.1; -;
DR TIGR; CCA00178; -;
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0015450; P:protein translocase activity; IEA.
DR GO; GO:0006886; P:intracellular protein transport; IEA.
DR GO; GO:0015628; P:type II protein (Sec) secretion system; IEA.
DR InterPro; IPR001005; Myd_DNA_binding.
DR InterPro; IPR005791; SecD.
DR Pfam; PF02355; SecD_SecF.
DR PRINTS; PR01755; SECYTRNLKASE.
DR TIGRPFAMS; TIGR00916; 2A0604801; 1.
DR TIGRPFAMS; TIGR01129; secD; 1.
DR PROSITE; PS00037; MYB_1; 1.
KW Complete proteome.
SQ SEQUENCE 1404 AA; 157493 MW; 804F644233BD1BD4 CRC64;

Query Match 86.7%; Score 39; DB 16; Length 1404;
Best Local Similarity 100.0%; Pred.No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 531 WVRWH 535

RESULT 18
QYU09 PRELIMINARY; PRT; 70 AA.
ID QYU09
AC QYU09
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

DE Neurotrophin receptor tyrosine kinase type 2 (Fragment).
GN NTRK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21656983; PubMed=11798182;
RA Stoilov P., Castren E., Stamm S.;
RT "Analysis of the Human TrkB Gene Genomic Organization Reveals Novel
RL TrkB Isoforms, Unusual Gene Length, and Splicing Mechanism.";
DR EMBL: AF410902; AAL67968.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00013; LRRNT; 1.
DR Receptor; Kinase.
KW NON_TER
FT SEQUENCE 70 AA; 7887 MW; D89B64C3D0CB7C3A CRC64;
SQ
Query Match 84.4%; Score 38; DB 4; Length 70;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db |:|
4 WIRWH 8
RESULT 19
Q9SKG5 PRELIMINARY; PRT; 178 AA.
AC Q9SKG5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RL libraries."
DR Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB066522; BAB62201.1; -.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00013; LRRNT; 1.
KW Hypothetical protein.
SQ SEQUENCE 178 AA; 20292 MW; A7B5CA4355FCD59F CRC64;
Query Match 84.4%; Score 38; DB 6; Length 178;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db |:|
4 WIRWH 8
RESULT 20
Q8WXJ6 PRELIMINARY; PRT; 537 AA.
ID Q8WXJ6
AC Q8WXJ6;

```

```

DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurotrophin receptor tyrosine kinase type 2 truncated isoform.
GN NTRK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21656983; PubMed=11798182;
RA Stoilov P., Castren E., Stamm S.;
RT "Analysis of the Human TrkB Gene Genomic Organization Reveals Novel
RL TrkB Isoforms, Unusual Gene Length, and Splicing Mechanism.";
DR EMBL; AF410900; AAL67966.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Kinase; Receptor.
SQ SEQUENCE 537 AA; 59166 MW; 5A8FA252A3871CC1 CRC64;
Query Match 84.4%; Score 38; DB 4; Length 537;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db |:|
4 WIRWH 8
RESULT 21
Q8WXJ5 PRELIMINARY; PRT; 553 AA.
AC Q8WXJ5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurotrophin receptor tyrosine kinase type 2 truncated isoform.
GN NTRK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21656983; PubMed=11798182;
RA Stoilov P., Castren E., Stamm S.;
RT "Analysis of the Human TrkB Gene Genomic Organization Reveals Novel
RL TrkB Isoforms, Unusual Gene Length, and Splicing Mechanism.";
DR EMBL; AF410901; AAL67967.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00047; IG; 1.

```

```
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00408; IGG2; 1.
DR SMART; SM00082; IGG2; 1.
DR SMART; SM00013; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Kinase; Receptor.
SQ SEQUENCE 553 AA; 60994 MW; BD98221B9EE1A6C1 CRC64;

Query Match 84.4%; Score 38; DB 4; Length 553;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVRWH 5
Db 4 WIRWH 8

RESULT 22
Q8WXJ7 PRELIMINARY; PRT; 838 AA.
AC Q8WXJ7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurotrophin receptor tyrosine kinase type 2 (BC 2.7.1.112) (Tyrosine-
DE protein kinase receptor).
GN NTRK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21656983; PubMed=11798182;
RA Stoilov P., Castren E., Stamm S.;
RT "Analysis of the Human TrkB Gene Genomic Organization Reveals Novel
RT TrkB Isoforms, Unusual Gene Length, and Splicing Mechanism.";
RL Biochem. Biophys. Res. Commun. 290:1054-1065(2002).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
CC RECEPTOR SUBFAMILY.
DR EMBL; AF410899; AAL67965.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. . .; IEA.
DR GO; GO:0008468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; RecepttyrkinII.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGG2; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00219; TyRK; 1.

DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 838 AA; 93825 MW; 130C95A9D8895432 CRC64;

Query Match 84.4%; Score 38; DB 4; Length 838;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVRWH 5
Db 4 WIRWH 8

RESULT 23
Q8GRG0 PRELIMINARY; PRT; 465 AA.
AC Q8GRG0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative amino acid transporter.
DE Arthrobacter nicotinovorans.
OS Arthrobacter pAO1.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=29320;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95115562; PubMed=7815950;
RA Græther-Beck S., Igloi G.L., Fust S., Schiltz E., Decker K.,
RA Brandsch R.;
RT "Structural analysis and molybdenum-dependent expression of the pAO1-
RT encoded nicotine dehydrogenase genes of Arthrobacter nicotinovorans.";
RL Mol. Microbiol. 13:929-936(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96172783; PubMed=8588735;
RA Menendez C., Igloi G., Hemminger H., Brandsch R.;
RT "A pAO1-encoded molybdopterin cofactor gene (moaA) of Arthrobacter
RT nicotinovorans: characterization and site-directed mutagenesis of the
RT encoded protein.";
RL Arch. Microbiol. 164:142-151(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97230479; PubMed=9073580;
RA Menendez C., Igloi G.L., Brandsch R.;
RT "IS1473, a putative insertion sequence identified in the plasmid pAO1
RT from Arthrobacter nicotinovorans: isolation, characterisation and
RT distribution among Arthrobacter species.";
RL Plasmid 37:35-41(1997).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=98088982; PubMed=9428706;
RA Menendez C., Otto A., Igloi G., Nick P., Brandsch R., Schubach B.,
RA Bottcher B., Brandsch R.;
RT "Molybdate-uptake genes and molybdopterin-biosynthesis genes on a
RT bacterial plasmid. Characterization of MoaA as a filament-forming
RT protein with adenosinetriphosphatase activity.";
RL Eur. J. Biochem. 250:524-531(1997).
RN [5]
RP SEQUENCE FROM N.A.
RA Schenk S., Hoelz A., Kraus B., Decker K.;
RT "Gene structure and properties of enzymes of the plasmid-encoded
RT nicotine catabolism of Arthrobacter nicotinovorans.";
RL J. Mol. Biol. 284:1323-1339(1999).
RN [6]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=21405725; PubMed=11514508;
RA Baitsch D., Sandu C., Brandsch R., Igloi G.L.;
RT "A gene cluster on pAO1 of Arthrobacter nicotinovorans involved in the
RT degradation of the plant alkaloid nicotine: cloning, purification and
RT characterization of 2,6-dihydroxypyridine 3-hydroxylase.";
RL J. Bacteriol. 183:5262-5267(2001).
RN [7]
RP SEQUENCE FROM N.A.
RA Igloi G.L., Brandsch R.;
RT "Sequence of the 165 kb Catabolic Plasmid pAO1 from Arthrobacter
RT nicotinovorans and Identification of a pAO1-dependent Nicotine Uptake
RT System.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ507836; CAD47971.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; F:Amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:Amino acid transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002293; AA/permease.
DR InterPro; IPR004841; Permease region.
DR Pfam; PF00324; aa_permeases; I.
KW Plasmid.
SQ SEQUENCE 465 AA; 48711 MW; 8ADBEF82E9D20F1D CRC64;

Query Match 82.2%; Score 37; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 454 WVRWF 459

RESULT 24
Q13041 PRELIMINARY; PRT; 591 AA.
AC Q13041;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P67.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hepatocytes;
RA Yeh P., Yew F.;
RT "Human p67 mRNA sequence.";
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U18247; AAA57339.1; -.
DR PIR; G01586; G01586.
DR InterPro; IPR002015; APC_proteasome.
DR InterPro; IPR008938; ARM.
DR Pfam; PF01851; PC_rep; 7.
SQ SEQUENCE 591 AA; 64587 MW; 4C819E01A924DBD4 CRC64;

Query Match 82.2%; Score 37; DB 4; Length 591;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 350 WVRWF 355

RESULT 25
Q8GEF9 PRELIMINARY; PRT; 99 AA.
ID Q8GEF9
AC Q8GEF9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
```

```
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Erwinia amylovora.
OG Plasmid pEA2.8.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL-5;
RX MEDLINE=22338281; PubMed=12450843;
RA McGhee G.C., Schnabel E.L., Maxson-Stein K., Jones B., Stromberg V.K.,
RA Lacy G.H., Jones A.L.;
RT "Relatedness of Chromosomal and Plasmid DNAs of Erwinia pyrifoliae and
RT Erwinia amylovora.";
RL Appl. Environ. Microbiol. 68:6182-6192(2002).
DR EMBL; AV123047; AAM94884.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 99 AA; 11282 MW; 9145B1CB833BA9E3 CRC64;

Query Match 80.0%; Score 36; DB 2; Length 99;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 21 WVRWF 26

RESULT 26
Q8LI79 PRELIMINARY; PRT; 115 AA.
ID Q8LI79;
AC Q8LI79;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OSJNB0093M23.10 protein.
GN OSJNB0093M23.10.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
RT clone:OSJNB0093M23.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003854; BAC06283.1; -.
DR Gramene; Q8LI79; -.
SQ SEQUENCE 115 AA; 12805 MW; BBFBF7F587093B55 CRC64;

Query Match 80.0%; Score 36; DB 10; Length 115;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 24 WVRWF 29

RESULT 27
Q7XKD7 PRELIMINARY; PRT; 177 AA.
ID Q7XKD7;
AC Q7XKD7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSJNB0017101.17 protein.
```

GN OSNBNB0017101.17.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL606456; CAFE05737.1; EF27848839DSB41B CRC64;
SQ SEQUENCE 177 AA; 20540 MW; EF27848839DSB41B CRC64;
Query Match 80.0%; Score 36; DB 10; Length 177;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db 42 WLRWH 46

RESULT 28
Q9S581
ID Q9S581 PRELIMINARY; PRT; 185 AA.
AC Q9S581;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE VR2 protein.
GN VR2 OR PA0616.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 15692 / PA01;
RA Nakayama K., Takashima K., Ishihara H., Shinomiya T., Kageyama M.,
RA Kanaya S., Ohnishi M., Murata T., Terawaki Y., Mori H., Hayashi T.;
RA "Genetic relationship between bacteriocins and bacteriophages."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AB030825; BA083153.1; -;
DR EMBL; AB004497; AAG04005.1; -;
DR PIR; T44538; T44538.
DR InterPro; IPR006531; Phage_P2_V.
DR Pfam; PF04717; phage_base_V; 1.
DR TIGRFAMs; TIGR01644; phage_P2_V; 1.
KW Complete proteome.
SQ SEQUENCE 185 AA; 19422 MW; 09E3DD38ED8A8D33 CRC64;
Query Match 80.0%; Score 36; DB 16; Length 185;

Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db 42 WLRWH 46

RESULT 29
Q88IE8
ID Q88IE8 PRELIMINARY; PRT; 200 AA.
AC Q88IE8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pycocin R2_PP, tail spike protein.
GN GPV OR PP3051.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
RA Uterback T., Rizzo M., Lee K., Koeck D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tsemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016785; AAN68659.1; -;
DR TIGR; PP3051; -;
DR InterPro; IPR006531; Phage_P2_V.
DR Pfam; PF04717; phage_base_V; 1.
KW Complete proteome.
SQ SEQUENCE 200 AA; 21248 MW; FEFD9F224482D298 CRC64;
Query Match 80.0%; Score 36; DB 16; Length 200;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db 43 WVRWH 47

RESULT 30
Q98BA4
ID Q98BA4 PRELIMINARY; PRT; 223 AA.
AC Q98BA4;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nodulation protein, Noe1.
GN MLL5661.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RL *Mesorhizobium loti*,"
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003007; BAB52068.1; -.
 DR InterPro; IPR006342; FkDM.
 DR TIGRFAMs; TIGR01444; fkdM_fam; 1.
 KW Complete proteome.
 SQ SEQUENCE 223 AA; 25091 MW; 9DC03946FA3A8FAC CRC64;

Query Match 80.0%; Score 36; DB 16; Length 223;
 Best Local Similarity 80.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWHF 5
 Db 12 WLRWH 16

RESULT 31
 OS3604 PRELIMINARY; PRT; 230 AA.
 AC OS3604;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein rv0059.
 GN RV0059 OR MT0030.02 OR MT0065.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaija F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544 (1998).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OshKosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL021428; CAAL6240.1; -.
 DR EMBL; AE006919; AAK44287.1; -.
 DR FIR; D70847; D70847.
 DR TIGR; MT0065; -.
 DR Tuberculist; rv0059; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 230 AA; 25580 MW; 1A109D34B16590A8 CRC64;

Query Match 80.0%; Score 36; DB 16; Length 230;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 Db 26 WIVWHF 31

RESULT 32

Q7U2Y3 PRELIMINARY; PRT; 230 AA.
 AC Q7U2Y3;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN MB0060.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis.";
 RL Proc. Natl Acad. Sci. U.S.A. 100:7877-7882(2003).
 DR EMBL; BX248334; CAD92922.1; -.
 KW Complete proteome.
 SQ SEQUENCE 230 AA; 25580 MW; 1A109D34B16590A8 CRC64;

Query Match 80.0%; Score 36; DB 16; Length 230;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 Db 26 WIVWHF 31

RESULT 33

Q9HYG3 PRELIMINARY; PRT; 262 AA.
 AC Q9HYG3;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Probable permease of ABC transporter.
 GN PA3443.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.I., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004765; AAG06831.1; -.
 DR FIR; H83214; H83214.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 KW Complete proteome.
 SQ SEQUENCE 262 AA; 28455 MW; CFB540F4F15BE25C CRC64;

Query Match 80.0%; Score 36; DB 16; Length 262;
 Best Local Similarity 80.0%; Pred. No. 2.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 Db 248 WLRWH 252
 : : : :
 : : : :
 : : : :

RESULT 34
 Q8XZQ5 PRELIMINARY; PRT; 264 AA.

AC Q8XZQ5; 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Putative aliphatic sulfonates transmembrane ABC transporter protein.
 DE protein.
 GN SSUC OR RS02862.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brotter P., Camus J.C., Cattolico L.,
 RA Chandler M., Cholsne N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Molzan A., Robert C., Saurin W., Schlex T.,
 RA Sigvier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RT Nature 415:497-502 (2002).
 RL EMBL; AL646064; CAD15042.1; --
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 KW Complete proteome.
 SQ SEQUENCE 264 AA; 29049 MW; BFE5AA19EB615D69 CRC64;

Query Match 80.0%; Score 36; DB 16; Length 264;
 Best Local Similarity 80.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 Db 253 WLRWH 257
 : : : :
 : : : :
 : : : :

RESULT 35
 Q8KZQ7 PRELIMINARY; PRT; 265 AA.

AC Q8KZQ7; 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE ABC-type transporter membrane permease component.
 GN DSUC.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DS1;
 RA Endoh T., Kasuga K., Horinouchi M., Yoshida T., Habe H., Nojiri H.,
 RA Omori T.;
 RA "Characterization and identification of genes essential for dimethyl

RT sulfide-utilization in pseudomonas putida strain DS1.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB086390; BAC00974.1; --
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 SQ SEQUENCE 265 AA; 28605 MW; 9AFFD2E536F23C4C CRC64;

Query Match 80.0%; Score 36; DB 2; Length 265;
 Best Local Similarity 80.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 Db 250 WLRWH 254
 : : : :
 : : : :
 : : : :

RESULT 36
 O85765 PRELIMINARY; PRT; 265 AA.

AC O85765;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE ABC transporter membrane subunit precursor.
 GN SSUC.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S-313;
 RX MEDLINE=99291059; PubMed=10361295;
 RA Vermeij P., Wietek C., Kahner A., Wuest T., Kertesz M.A.;
 RT "Genetic organization of sulphur-controlled aryl desulphonation in
 Pseudomonas putida S-313.";
 RL Mol. Microbiol. 32:913-926 (1999).
 DR EMBL; AF075709; AAC31906.1; --
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 KW SIGNAL.
 FT SIGNAL.
 SQ SEQUENCE 265 AA; 28595 MW; 576417A6BDECA6B CRC64;

Query Match 80.0%; Score 36; DB 2; Length 265;
 Best Local Similarity 80.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
 Db 250 WLRWH 254
 : : : :
 : : : :
 : : : :

RESULT 37
 Q8ZB03 PRELIMINARY; PRT; 265 AA.

AC Q8ZB03; 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Putative aliphatic sulfonates transport permease protein (Putative
 DE transport system permease protein of aliphatic sulfonates ABC
 DE transporter).
 GN SSUC OR YP03626 OR Y0243.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.

```

OX NCBI_TaxID=632;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN
DR EMBL; AJ414158; CAC93095.1; -.
DR EMBL; AE013624; AAM83837.1; -.
DR PIR; AC0441; AC0441.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp; 1.
DR Hypothetical protein; Complete proteome.
RW
SQ SEQUENCE 265 AA; 29011 MW; 26296549CFD42BDA CRC64;

Query Match 80.0%; Score 36; DB 16; Length 265;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
Db 250 WLRWH 254

RESULT 38
Q88R94 PRELIMINARY; PRT; 265 AA.
ID Q88R94
AC Q88R94;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sulfonate ABC transporter, permease protein SsUC.
GN SSUC OR P20239.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapflee E., Scanlan D., Tran K., Moazzes A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).

```

```

DR EMBL; AE016774; AAN65871.1; -.
DR TIGR; P20239; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp; 1.
DR Complete proteome.
RW
SQ SEQUENCE 265 AA; 28665 MW; PDD7354A277CSA27 CRC64;

Query Match 80.0%; Score 36; DB 16; Length 265;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
Db 250 WLRWH 254

RESULT 39
Q8U649 PRELIMINARY; PRT; 266 AA.
ID Q8U649
AC Q8U649;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ABC transporter, membrane spanning protein.
GN NRTB OR ATU6086 OR AGR_PII_164.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse K., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Liu F.,
RA Houmel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
RN
DR EMBL; AE009427; AAL46322.1; -.
DR EMBL; AE007935; AAK91046.1; -.
DR PIR; AD3238; AD3238.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp; 1.
DR Complete proteome.
RW
SQ SEQUENCE 266 AA; 28817 MW; 814ACD84FCDB8D73 CRC64;

```

```

Query Match      80.0%; Score 36; DB 16; Length 266;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 257 WLRWH 261

RESULT 40
Q9RS74          PRELIMINARY; PRT; 384 AA.
AC Q9RS74;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein DR2253.
GN DR2253.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Hatt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE002058; AAF11801.1; -.
DR PIR; E75295; E75295.
DR TIGR; DR2253; -.
DR InterPro; IPR007299; DUF405.
DR InterPro; IPR007349; DUF418.
DR Pfam; PF04171; DUF405; 1.
DR Pfam; PF04235; DUF418; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 384 AA; 42270 MW; B01C4515B4D0627 CRC64;

Query Match      80.0%; Score 36; DB 16; Length 394;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 362 WLRWH 366

RESULT 41
Q8YMN8          PRELIMINARY; PRT; 472 AA.
AC Q8YMN8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein A114895.
GN A114895.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

```

```

RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003598; BAB76594.1; -.
DR PIR; AG2417; AG2417.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR004294; RPE65.
DR InterPro; IPR001680; WD40.
DR Pfam; PF03055; RPE65; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 472 AA; 52893 MW; E7863260EC1E5616 CRC64;

Query Match      80.0%; Score 36; DB 16; Length 472;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 6
DB 285 WYQWH 290

RESULT 42
Q8PFC2          PRELIMINARY; PRT; 475 AA.
AC Q8PFC2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XAC4061.
GN XAC4061.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A., P.,
RA Ciccarotti G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE012054; AAM38896.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 475 AA; 51733 MW; 064D5BEE2363C41F CRC64;

Query Match      80.0%; Score 36; DB 16; Length 475;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
DB 237 WLRWH 241

```

```

RESULT 43
Q88HM3 ID Q88HM3 PRELIMINARY; PRT; 497 AA.
AC Q88HM3
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN PF3331.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzaz A.,
RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808 (2002).
DR TIGR; PF3331; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 497 AA; 55110 MW; 3C0EC5CCBC652AEF CRC64;

Query Match 80.0%; Score 36; DB 16; Length 497;
Best Local Similarity 80.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
Db 251 WLRWH 255

RESULT 44
Q82Q09 ID Q82Q09 PRELIMINARY; PRT; 699 AA.
AC Q82Q09
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN SAV396.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT

```

```

RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531 (2003).
DR EMBL; AP005022; BAC68105.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 699 AA; 76830 MW; 6098AD18A0C68174 CRC64;

Query Match 80.0%; Score 36; DB 16; Length 699;
Best Local Similarity 66.7%; Pred. No. 6.6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 6
Db 323 WVSWHY 328

RESULT 45
Q8ZQC3 ID Q8ZQC3 PRELIMINARY; PRT; 754 AA.
AC Q8ZQC3
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative recombinant protein.
GN YCAI OR STM0983.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SGSC1412 / ATCC 700720;
MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
RX EMBL; AE008742; AAL19917.1; -.
GO GO; GO:0016021; C:integral to membrane; IEA.
GO GO; GO:0030420; P:establishment of competence for transformation; IEA.
DR InterPro; IPR001279; Clactamase-like.
DR InterPro; IPR004477; ComEC N-term.
DR InterPro; IPR004797; ComEC Rec2.
DR Pfam; PF03772; Competence; 1.
DR Pfam; PF00753; lactamase B; 1.
DR TIGRFAMs; TIGR00360; ComEC N-term; 1.
DR TIGRFAMs; TIGR00361; ComEC Rec2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 754 AA; 84954 MW; 711A95D282271358 CRC64;

Query Match 80.0%; Score 36; DB 16; Length 754;
Best Local Similarity 80.0%; Pred. No. 7.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5
Db 542 WLRWH 546

RESULT 46
Q82802 ID Q82802 PRELIMINARY; PRT; 754 AA.
AC Q82802
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative competence-related protein.
GN STY0984 OR YCAI OR TI951.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham P., Brooks K., Chillingworth T., Connerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jørgensen K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 DR ENBL; AL627268; CAP05383.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0030420; P:stabilization of competence for transformation; IEA.
 DR InterPro; IPR001279; Bactamase-like.
 DR InterPro; IPR004477; ComEC N-term.
 DR Pfam; PF03772; Competence_1.
 DR Pfam; PF00753; lactamase B, 1.
 DR TIGRams; TIGR00360; ComEC N-term; 1.
 DR TIGRams; TIGR00361; ComEC Rec2; 1.
 SW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 754 AA; 84920 MW; 6697C033FB29576E CRC64;
 Query Match 80.08; Score 36; DB 16; Length 754;
 Best Local Similarity 80.08; Pred. No. 7.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db 542 WLRWH 546
 ID Q8C771 PRELIMINARY; PRT; 774 AA.
 AC Q8C771;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN D930015E06RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=22354685; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR ENBL; AK052409; BAC34977.1; -;
 DR MGD; MGI:2443399; D930015E06RIK.

KW Hypothetical protein.
 FT NON TER 774 774
 SQ SEQUENCE 774 AA; 86028 MW; 6AD44F81DF68FCAD CRC64;
 Query Match 80.0%; Score 36; DB 11; Length 774;
 Best Local Similarity 66.7%; Pred. No. 7.3e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WVRWH 6
 Db 760 WIFWH 765
 ID Q9VFX7 PRELIMINARY; PRT; 1169 AA.
 AC Q9VFX7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE SU(VAR)3-7 protein.
 DE SU(VAR)3-7 OR SU OR CG8599.
 OS Drosophila melanogaster (fruit fly).
 GN Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.P.,
 RA Sutton G.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Feiriza S., Fleischmann W.,
 RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Moulton G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR ENBL; AE003699; AAF54918.1; -;
 DR FlyBase; FBgn0003598; Su(Var)3-7.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR004210; BESS motif.

DR InterPro; IPR007087; Znf_C2H2.

DR Pfam; PF02944; BESS; 1.

DR SMART; SM00355; Znf_C2H2; 1.

DR PROSITE; PS00028; ZINC FINGER C2H2.1; 1.

SQ SEQUENCE 1169 AA; 131236 MW; 5F408B4E4564E8EC CRC64;

Query Match 80.0%; Score 36; DB 5; Length 1169;

Best Local Similarity 80.0%; Pred. No. 1.1e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5

Db 219 WLRWH 223

RESULT 49

Q8WT43

ID Q8WT43 PRELIMINARY; PRT; 1175 AA.

AC Q8WT43;

DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Hypothetical protein (Ubiquitin c-terminal hydrolase).

GN ZK328.1 OR CYK-3

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

[1] NCB

PP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for

investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).

FN [2]

PP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Favello A.;

RT "The sequence of C. elegans cosmid ZK328.";

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

FN [3]

PP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RT "Direct Submission.";

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

FN [4]

PP SEQUENCE FROM N.A.

RA Kaitna S., Schnabel H., Schnabel R., Hyman A.A., Glotzer M.;

RT "An ubiquitin C-terminal hydrolase is required to maintain osmotic

balance and execute actin-dependent processes in the early

Caenorhabditis elegans embryo.";

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; U50193; AAL32262.1; -

DR EMBL; AF469173; AAL79016.1; -

DR WormPep; ZK328.1a; CE30061.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.

DR GO; GO:0006512; P:ubiquitin cycle; IEA.

DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.

DR InterPro; IPR006615; DUSP.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR001394; Peptidase_C19.

DR Pfam; PF00036; ehand; 3.

DR Pfam; PF00443; UCH; 1.

DR SMART; SM00695; DUSP; 1.

DR SMART; SM00054; EFh; 2.

DR PROSITE; PS00018; EF_HAND; 2.

DR PROSITE; PS50235; UCH_2_3; 1.

KW Hypothetical protein; Hydrolase.

SQ SEQUENCE 1175 AA; 133832 MW; 734DE99C5ABF77A7 CRC64;

Query Match 80.0%; Score 36; DB 5; Length 1175;

Best Local Similarity 80.0%; Pred. No. 1.1e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5

Db 391 WLRWH 395

RESULT 50

Q8WT44

ID Q8WT44 PRELIMINARY; PRT; 1178 AA.

AC Q8WT44;

DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN ZK328.1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

[1] NCB

PP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for

investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).

FN [2]

PP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Favello A.;

RT "The sequence of C. elegans cosmid ZK328.";

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

FN [3]

PP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RT "Direct Submission.";

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; U50193; AAL32263.1; -

DR WormPep; ZK328.1b; CE30062.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.

DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.

DR GO; GO:0006512; P:ubiquitin cycle; IEA.

DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.

DR InterPro; IPR006615; DUSP.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR001394; Peptidase_C19.

DR Pfam; PF00036; ehand; 3.

DR Pfam; PF00443; UCH; 1.

DR SMART; SM00695; DUSP; 1.

DR SMART; SM00054; EFh; 2.

DR PROSITE; PS00018; EF_HAND; 2.

DR PROSITE; PS50235; UCH_2_3; 1.

KW Hypothetical protein.

SQ SEQUENCE 1178 AA; 134122 MW; 34F1FBA55DF4B6CA CRC64;

Query Match 80.0%; Score 36; DB 5; Length 1178;

Best Local Similarity 80.0%; Pred. No. 1.1e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRWH 5

Db 391 WLRWH 395

Search completed: June 10, 2004, 10:50:35
Job time : 49 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:36:04 ; Search time 48 Seconds
(without alignments)
35.318 Million cell updates/sec

Title: US-09-912-414-9

Perfect score: 31

Sequence: 1 WXXWXF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	90.3	15	2 AAY30351	Aay30351 Epitope d
2	28	90.3	15	5 AAE19239	Aae19239 Streptoco
3	28	90.3	47	5 AAE19247	Aae19247 Tri-pepti
4	28	90.3	57	4 AAU49566	Aau49566 Propionib
5	28	90.3	57	6 ABM46085	Abm46085 Propionib
6	28	90.3	64	5 AAE19246	Aae19246 Bi-peptid
7	28	90.3	66	5 AAE19245	Aae19245 Bi-peptid
8	28	90.3	67	5 ABB54348	Abb54348 Lactococc
9	28	90.3	95	4 AAU64053	Aau64053 Propionib
10	28	90.3	95	6 ABM60572	Abm60572 Propionib
11	28	90.3	57	6 ABP78009	Abp78009 N. gonorr
12	27	87.1	9	5 AAE26751	Aae26751 Fibrin bi
13	27	87.1	15	5 AAE26733	Aae26733 Fibrin bi
14	27	87.1	18	7 ADC99282	Adc99282 Cancer-re
15	27	87.1	19	5 AAE26744	Aae26744 Slow diss
16	27	87.1	19	5 AAE26745	Aae26745 Slow diss
17	27	87.1	19	5 AAE26743	Aae26743 Slow diss
18	27	87.1	21	4 ABG28607	Abg28607 Novel hum
19	27	87.1	23	5 ABG69143	Abg69143 H11 ligan
20	27	87.1	34	3 AAB34467	Aab34467 Human sec
21	27	87.1	52	4 AAM20383	Aam20383 Peptide #
22	27	87.1	52	4 AAB41043	Abb41043 Peptide #
23	27	87.1	52	4 AAM34818	Aam34818 Peptide #
24	27	87.1	52	4 ABB25116	Abb25116 Protein #
25	27	87.1	52	4 AAM74702	Aam74702 Human bon

26	27	87.1	52	4 AAM61900	Aam61900 Human bra
27	27	87.1	52	4 ABG56486	Abg56486 Human liv
28	27	87.1	52	5 ABG44513	Abg44513 Human pep
29	27	87.1	53	5 AAE21233	Aae21233 Human gen
30	27	87.1	53	5 ABG64892	Abg64892 Human alb
31	27	87.1	56	4 ABB17168	Abb17168 Human ner
32	27	87.1	103	5 ABB89068	Abb89068 Human pol
33	27	87.1	109	4 AAB66406	Aab66406 Human fab
34	27	87.1	120	4 AAM88588	Aam88588 Human imm
35	27	87.1	129	5 ABG93128	Abg93128 S. cerevi
36	27	87.1	145	7 ADB68030	Adb68030 Robacco F
37	27	87.1	146	5 ABG93116	Abg93116 S. cerevi
38	27	87.1	148	5 AAU95372	Aau95372 Human cal
39	27	87.1	155	7 ADC01647	Adc01647 Enteroha
40	27	87.1	163	4 AAU47693	Aau47693 Propionib
41	27	87.1	163	6 ABM44212	Abm44212 Propionib
42	27	87.1	164	6 ADB11318	Adb11318 Allolococ
43	27	87.1	167	6 ADB11316	Adb11316 Allolococ
44	27	87.1	169	3 AAY74781	Aay74781 Neisseria
45	27	87.1	186	4 ABB17902	Abb17902 Human ner
46	27	87.1	186	5 ABB49316	Abb49316 Listeria
47	27	87.1	214	3 AAY74779	Aay74779 Neisseria
48	27	87.1	216	3 AAY74778	Aay74778 Neisseria
49	27	87.1	216	3 AAB13554	Abp13554 Streptom
50	27	87.1	216	6 ABP80786	Abp80786 N. gonorr
51	27	87.1	217	3 AAY74780	Aay74780 Neisseria
52	27	87.1	221	3 AAY74783	Aay74783 Neisseria
53	27	87.1	221	3 AAY74782	Aay74782 Neisseria
54	27	87.1	241	3 AAG36488	Aag36488 Arabidops
55	27	87.1	245	5 ABP57756	Abp57756 E. coli L
56	27	87.1	260	4 ABG17243	Abg17243 Novel hum
57	27	87.1	293	3 AAG36487	Aag36487 Arabidops
58	27	87.1	294	3 AAG36486	Aag36486 Arabidops
59	27	87.1	294	6 ABG71364	Abg71364 Potato st
60	27	87.1	306	6 ABM68614	Abm68614 Photorhab
61	27	87.1	312	5 ABP73273	Abp73273 Candida a
62	27	87.1	360	6 ABU25784	Abu25784 Protein e
63	27	87.1	416	2 AAR97658	Aar97658 Zebrafish
64	27	87.1	416	2 AAM61485	Aam61485 Zebrafish
65	27	87.1	416	2 AAW94475	Aaw94475 Zebrafish
66	27	87.1	416	2 AAY05862	Aay05862 Zebrafish
67	27	87.1	416	2 AAW97771	Aaw97771 Zebrafish
68	27	87.1	416	2 AAY05518	Aay05518 Zebrafish
69	27	87.1	416	3 AAY96251	Aay96251 Zebrafish
70	27	87.1	416	3 AAY70684	Aay70684 Zebrafish
71	27	87.1	416	3 AAY95289	Aay95289 Zebrafish
72	27	87.1	416	3 AAY95980	Aay95980 Zebrafish
73	27	87.1	416	4 AAB84677	Aab84677 Amino aci
74	27	87.1	416	4 AAB60268	Aab60268 Zebrafish
75	27	87.1	416	4 AAE04706	Aae04706 Zebrafish
76	27	87.1	416	4 AAG65751	Aag65751 Zebrafish
77	27	87.1	416	5 AAE14300	Aae14300 Zebrafish
78	27	87.1	416	5 AAO20926	Aao20926 Zebrafish
79	27	87.1	416	5 AAU99486	Aau99486 Zebrafish
80	27	87.1	425	2 AAR77338	Aar77338 Chicken S
81	27	87.1	425	2 AAW94468	Aaw94468 Chicken S
82	27	87.1	425	2 AAY05854	Aay05854 Chicken S
83	27	87.1	425	2 AAW97765	Aaw97765 Chicken S
84	27	87.1	425	2 AAY05510	Aay05510 Chicken S
85	27	87.1	425	3 AAY96243	Aay96243 Partial C
86	27	87.1	425	3 AAY70676	Aay70676 Chicken S
87	27	87.1	425	3 AAY95281	Aay95281 Chicken S
88	27	87.1	425	3 AAY95972	Aay95972 Chicken S
89	27	87.1	425	4 AAB84669	Aab84669 Amino aci
90	27	87.1	425	4 AAB60260	Aab60260 Chicken S
91	27	87.1	425	4 AAE04682	Aae04682 Chicken S
92	27	87.1	425	4 AAB05372	Aab05372 Chicken S
93	27	87.1	425	4 AAB85733	Aab85733 Chicken S
94	27	87.1	425	4 AAG65743	Aag65743 Chicken S
95	27	87.1	425	4 AAB31217	Aab31217 Amino aci
96	27	87.1	425	5 ABB79132	Abb79132 Chicken S
97	27	87.1	425	5 AAE14292	Aae14292 Chicken S
98	27	87.1	425	5 AAO20918	Aao20918 Chicken S

99	27	87.1	425	5	AAU99478	Rau99478	Chicken s	172	26	83.9	68	4	AAU44646	Aau44646	Propionib
100	27	87.1	425	6	ADA26251	Ada26251	Chicken S	173	26	83.9	68	6	ABW41165	Abm41165	Propionib
101	27	87.1	425	7	ADB25255	Adw25255	Chicken S	174	26	83.9	69	4	AGG76062	Agp76062	Human col
102	27	87.1	425	8	ABW00863	Abw00863	Chicken s	175	26	83.9	74	5	ABP04814	Abp04814	Human ORF
103	27	87.1	425	9	ADD71378	Adp71378	Chicken s	176	26	83.9	77	4	AAU66827	Aau66827	Novel hum
104	27	87.1	449	5	ABP66103	Abp66103	BiFidobac	177	26	83.9	77	7	ADB60161	Adb60161	Connectiv
105	27	87.1	520	4	AGG90739	Agg90739	C glutami	178	26	83.9	78	2	AAW78321	Aaw78321	Fragment
106	27	87.1	579	6	ABU17611	Abu17611	Protein e	179	26	83.9	82	4	ABE11005	Abb11005	Human sec
107	27	87.1	583	6	ADA35135	Ada35135	Acinetoba	180	26	83.9	83	5	ABB99041	Abb99041	Serine/th
108	27	87.1	700	6	ABU49712	Abu49712	Protein e	181	26	83.9	84	3	AAW34708	Aab34708	Human sec
109	27	87.1	709	4	ABE61518	Abb61518	Drosophil	182	26	83.9	85	4	AAU59296	Aau59296	Propionib
110	27	87.1	768	5	AAW27684	Aar27684	Human bet	183	26	83.9	85	4	ABG24922	Abg24922	Novel hum
111	27	87.1	769	5	ABG96351	Abg96351	Human ova	184	26	83.9	85	6	ABM55815	Abm55815	Propionib
112	27	87.1	769	6	ABU56723	Abu56723	Lung canc	185	26	83.9	86	4	AAW88451	Aam88451	Human imm
113	27	87.1	769	7	ADB80492	Adb80492	Ovarian c	186	26	83.9	87	3	AAW56431	Aab56431	Human pro
114	27	87.1	769	8	ADB86616	Adb86616	Human int	187	26	83.9	90	6	ABP75777	Abp75777	Human sec
115	27	87.1	943	6	AAW20588	Abu20588	Protein e	188	26	83.9	93	4	AAW90556	Aag90556	Human imm
116	27	87.1	1068	4	AGG67244	Agg67244	Amino aci	189	26	83.9	100	3	AAW38449	Aag38449	Arabidops
117	27	87.1	4472	2	AAW97245	Agp7245	Virulence	190	26	83.9	105	3	AAW58425	Aag58425	Arabidops
118	27	87.1	19938	6	ABP76681	Abp76681	Streptomy	191	26	83.9	105	4	AAU14416	Aau14416	Human nov
119	27	87.1	19938	6	ABP76678	Abp76678	Streptomy	192	26	83.9	108	3	AAW42837	Aab42837	Human ORF
120	26	83.9	6	3	AAW01508	Abp01508	Peptide w	193	26	83.9	108	6	ABR54906	AbR54906	Light cha
121	26	83.9	6	3	AAW01505	Abp01505	Peptide w	194	26	83.9	109	4	AAW93527	Aam93527	Human pol
122	26	83.9	6	3	AAW01506	Abp01506	Peptide w	195	26	83.9	110	3	AAW00830	Aag00830	Human sec
123	26	83.9	6	6	ABR45426	AbR45426	Staphyloc	196	26	83.9	112	3	AAW01617	Agw01617	Human sec
124	26	83.9	6	6	ABR45594	AbR45594	Staphyloc	197	26	83.9	112	5	ABP39843	Abp39843	Staphyloc
125	26	83.9	6	6	ABR45314	AbR45314	Staphyloc	198	26	83.9	113	6	ABG18237	Abg18237	Novel hum
126	26	83.9	6	6	ABR45481	AbR45481	Staphyloc	199	26	83.9	113	6	ADA55310	Ada55310	Human pro
127	26	83.9	6	6	ABR45369	AbR45369	Staphyloc	200	26	83.9	114	4	AAW66330	Aar66330	Human imm
128	26	83.9	6	6	ABR45425	AbR45425	Staphyloc	201	26	83.9	115	7	ADB65752	AdB65752	Human pro
129	26	83.9	6	6	ABR45593	AbR45593	Staphyloc	202	26	83.9	116	2	AAW66341	Aar66341	Human imm
130	26	83.9	6	6	ABR45313	AbR45313	Staphyloc	203	26	83.9	118	2	AAW66314	Aar66314	Human imm
131	26	83.9	6	6	ABR45370	AbR45370	Staphyloc	204	26	83.9	119	4	AAW94671	Aar94671	Human pro
132	26	83.9	6	6	ABR45482	AbR45482	Staphyloc	205	26	83.9	120	5	ABW55413	AbW55413	Lactococc
133	26	83.9	9	5	AAE26775	Aae26775	Fibrin bi	206	26	83.9	121	6	ADB10001	AdB10001	Alloiococ
134	26	83.9	15	3	AAW65508	Aay65508	Oestrogen	207	26	83.9	123	3	AAW42875	Aab42875	Human ORF
135	26	83.9	15	5	AAE26759	Aae26759	Fibrin bi	208	26	83.9	124	5	ABP38085	Abp38085	Staphyloc
136	26	83.9	15	5	AAU86245	Aau86245	Oestrogen	209	26	83.9	127	3	AAW27726	Aab27726	Human sec
137	26	83.9	15	5	ABW99042	Abw99042	Serine/th	210	26	83.9	128	5	ABP43136	Abp43136	Human ova
138	26	83.9	30	4	ABW50643	AbW50643	Human sec	211	26	83.9	132	4	ABW32752	AbW32752	Peptide #
139	26	83.9	30	6	ABO44900	AbO44900	Novel hum	212	26	83.9	132	4	AAW26214	Aam26214	Peptide #
140	26	83.9	30	7	ABO26380	AbO26380	Protein a	213	26	83.9	132	4	ABW18234	AbW18234	Protein #
141	26	83.9	33	6	AAE36631	Aae36631	Staphyloc	214	26	83.9	132	5	ABW35584	AbW35584	Human pep
142	26	83.9	36	4	ABW40893	AbW40893	Peptide #	215	26	83.9	134	4	ABW14102	AbW14102	Novel hum
143	26	83.9	36	4	AAW34663	Aam34663	Peptide #	216	26	83.9	135	4	ABW17562	AbW17562	Human ner
144	26	83.9	36	4	AAW74549	Aam74549	Human bon	217	26	83.9	137	4	AAW75484	Aag75484	Human col
145	26	83.9	36	4	AAW61753	Aam61753	Human bra	218	26	83.9	137	6	ABP75551	AbP75551	Human sec
146	26	83.9	36	4	ABG56337	AbG56337	Human liv	219	26	83.9	138	3	AAW42107	AbA42107	Human ORF
147	26	83.9	36	6	ABJ19233	AbJ19233	Staphyloc	220	26	83.9	141	5	ABP09777	AbP09777	Human ORF
148	26	83.9	41	4	AAE011216	Aae011216	Human gen	221	26	83.9	141	5	AAW87493	AaW87493	Novel cen
149	26	83.9	41	4	AAE011178	Aae011178	Human gen	222	26	83.9	145	4	AAW17192	AaW17192	Human sec
150	26	83.9	41	4	AAE011217	Aae011217	Human gen	223	26	83.9	147	5	AAW17226	AaW17226	Human sec
151	26	83.9	41	5	ABG64048	AbG64048	Human alb	224	26	83.9	147	5	ABG64768	AbG64768	Human alb
152	26	83.9	41	5	ABG64046	AbG64046	Human alb	225	26	83.9	147	5	ABG64767	AbG64767	Human alb
153	26	83.9	41	5	ABG64047	AbG64047	Human alb	226	26	83.9	149	5	ABU61012	AbU61012	Lung spec
154	26	83.9	43	2	AAW55277	Aaw55277	CD44 Exon	227	26	83.9	157	6	ABU29479	AbU29479	Protein e
155	26	83.9	43	2	AAW47930	AaW47930	Polypepti	228	26	83.9	159	2	AAW44125	AaW44125	Streptoco
156	26	83.9	44	2	AAW12655	Aay12655	Human 5'	229	26	83.9	160	3	AAW18297	AaW18297	Arabidops
157	26	83.9	46	2	AAW48521	Aay48521	Human bre	230	26	83.9	161	3	AAW38448	AaW38448	Arabidops
158	26	83.9	47	4	AAW88166	Aam88166	Human imm	231	26	83.9	163	3	AAW18296	AaW18296	Arabidops
159	26	83.9	48	6	ABU07201	AbU07201	Maize SSI	232	26	83.9	164	7	ADE59161	AdE59161	Human Pro
160	26	83.9	48	6	ABU07199	AbU07199	Maize SSI	233	26	83.9	165	4	AAW88604	AaW88604	Human pol
161	26	83.9	50	5	ABP04232	AbP04232	Human ORF	234	26	83.9	165	5	ABW89697	AbW89697	Human hyd
162	26	83.9	51	2	AAW38387	Aay38387	Human sec	235	26	83.9	165	5	ABW06127	AbW06127	Human NS
163	26	83.9	52	3	AAW76057	Aay76057	Human ski	236	26	83.9	165	6	ABU00301	AbU00301	Human nov
164	26	83.9	52	4	AAW55996	Aab55996	Skin cell	237	26	83.9	166	3	AAW18295	AaW18295	Arabidops
165	26	83.9	52	5	ABW72196	AbW72196	Human pro	238	26	83.9	167	3	AAW38447	AaW38447	Arabidops
166	26	83.9	60	4	AAW85125	Aam85125	Human imm	239	26	83.9	168	4	ABG19869	AbG19869	Novel hum
167	26	83.9	60	4	AAW49249	AaW49249	Propionib	240	26	83.9	169	6	ABP77127	AbP77127	N. gonorr
168	26	83.9	60	6	ABW45768	AbW45768	Propionib	241	26	83.9	173	4	AAU59322	Aau59322	Propionib
169	26	83.9	65	5	ABP07322	AbP07322	Human ORF	242	26	83.9	173	6	ABW55841	AbW55841	Propionib
170	26	83.9	67	5	ABP00507	AbP00507	Human ORF	243	26	83.9	175	4	ABW03211	AbW03211	Human mus
171	26	83.9	67	5	ABP31356	AbP31356	Human ORF	244	26	83.9	175	4	AAW42371	AaW42371	Human pol

245	26	83.9	175	6	ABU12505	Abu12505 Novel hum	318	26	83.9	295	3	AAG52323	Aag52323 Arabidops
246	26	83.9	177	7	ADC97284	E. faeciu	319	26	83.9	301	5	ABBA48598	Abba48598 Listeria
247	26	83.9	181	4	ABG28127	Novel hum	320	26	83.9	303	5	ABG27536	Abg27536 Arabidops
248	26	83.9	183	4	ABG25122	Novel hum	321	26	83.9	306	6	ADA36890	Ada36890 Acinetoba
249	26	83.9	187	4	ABG70657	Drosophil	322	26	83.9	310	2	AAZ21590	Aaz21590 Human sec
250	26	83.9	193	4	ABG71944	Human olf	323	26	83.9	310	7	ADC38838	Adc38838 Human sec
251	26	83.9	193	6	ADA34747	Acinetoba	324	26	83.9	311	3	AAG20705	Aag20705 Arabidops
252	26	83.9	197	3	AAG59004	Arabidops	325	26	83.9	314	5	ABBA48389	Abba48389 Listeria
253	26	83.9	197	5	ABB06830	Human ngp	326	26	83.9	314	6	ABUJ33088	AbuJ33088 Protein e
254	26	83.9	199	3	AAG36625	Arabidops	327	26	83.9	314	6	ABU17459	Abu17459 Protein e
255	26	83.9	200	3	AAG59003	Arabidops	328	26	83.9	315	3	AAG05942	Aag05942 Arabidops
256	26	83.9	200	4	ABM50640	Human sec	329	26	83.9	315	6	ABU29298	Abu29298 Protein e
257	26	83.9	200	6	ABM15921	Mycobacte	330	26	83.9	315	6	ADB07220	Adb07220 Alloiooc
258	26	83.9	200	6	ABO44897	Novel hum	331	26	83.9	316	4	ABG17054	Abg17054 Novel hum
259	26	83.9	200	6	ADB09999	Novel hum	332	26	83.9	316	5	AAM49496	Aam49496 Lactobaci
260	26	83.9	200	7	ABO26377	Protein a	333	26	83.9	317	2	RAY23764	Ray23764 A carboxy
261	26	83.9	205	3	AAG58424	Arabidops	334	26	83.9	317	2	AAZ39338	Aaz39338 Carboxyme
262	26	83.9	206	4	ABG04155	Novel hum	335	26	83.9	317	3	AAZ56814	Aaz56814 T. mariti
263	26	83.9	208	3	AAG22447	Arabidops	336	26	83.9	317	3	AAG05941	Aag05941 Arabidops
264	26	83.9	209	3	ABG11230	Arabidops	337	26	83.9	317	4	AAE08543	Aae08543 Thermotog
265	26	83.9	210	4	ABB67271	Drosophil	338	26	83.9	317	6	ABP56759	Abp56759 T. mariti
266	26	83.9	211	6	ABM67924	Phototrab	339	26	83.9	318	6	ABU42411	Abu42411 Protein e
267	26	83.9	214	3	AAG36623	Arabidops	340	26	83.9	318	6	ABU42887	Abu42887 Protein e
268	26	83.9	217	2	AAW69658	Human euk	341	26	83.9	318	6	ABU43921	Abu43921 Protein e
269	26	83.9	217	2	AAW69655	Human euk	342	26	83.9	319	3	AAG11228	Aag11228 Arabidops
270	26	83.9	217	2	AAW69656	Human euk	343	26	83.9	319	3	AAG12074	Aag12074 Arabidops
271	26	83.9	217	2	AAW69657	Human euk	344	26	83.9	321	5	ABP38349	Abp38349 Staphyloc
272	26	83.9	217	2	AAW94251	Murine eu	345	26	83.9	322	5	ABP38349	Abp38349 Staphyloc
273	26	83.9	217	3	AAW78505	Human gen	346	26	83.9	322	6	ABP57980	Abp57980 Outer mem
274	26	83.9	217	3	AAW78503	Murine ge	347	26	83.9	322	6	ABP38883	Abp38883 Surface e
275	26	83.9	217	4	ABG14213	Novel hum	348	26	83.9	322	6	ABM73020	Abm73020 Staphyloc
276	26	83.9	217	7	ADD48296	Rat Prote	349	26	83.9	324	3	AAG50316	Aag50316 Arabidops
277	26	83.9	217	7	ADD48298	Human Pro	350	26	83.9	324	7	ADC96518	Adc96518 E. faeciu
278	26	83.9	218	4	ABG27472	Novel hum	351	26	83.9	325	3	AAG48292	Aag48292 Arabidops
279	26	83.9	223	4	AAW76835	Corynebac	352	26	83.9	325	3	AAG44074	Aag44074 Arabidops
280	26	83.9	223	4	ABB62315	Drosophil	353	26	83.9	325	3	AAG09462	Aag09462 Arabidops
281	26	83.9	223	4	AAW91884	C glutami	354	26	83.9	325	7	ADB95074	Adb95074 A. thalia
282	26	83.9	224	2	AAW55278	Ar55278 HIV2 (gp32	355	26	83.9	333	4	AAG91017	Aag91017 C glutami
283	26	83.9	224	2	AAW81138	Mycobacte	356	26	83.9	333	7	ADD24923	Add24923 Ralstonia
284	26	83.9	229	2	AAW52985	Novel hum	357	26	83.9	335	5	ABB92415	Abb92415 Herbicida
285	26	83.9	238	4	ABB68592	Drosophil	358	26	83.9	338	4	AAU35351	Aau35351 Enterococ
286	26	83.9	241	4	ABB70112	Drosophil	359	26	83.9	338	4	AAU35351	Aau35351 Enterococ
287	26	83.9	242	3	AAG05943	Arabidops	360	26	83.9	340	6	ABU29420	Abu29420 Protein e
288	26	83.9	242	7	ADC31813	Human nov	361	26	83.9	346	6	AAU34331	Aau34331 Protein e
289	26	83.9	251	3	ABG40843	Zea mays	362	26	83.9	348	3	AAU29338	Aau29338 Novel mar
290	26	83.9	254	4	ABB69619	Drosophil	363	26	83.9	348	4	ABB64322	Abb64322 Drosophil
291	26	83.9	254	6	ABU23970	Protein e	364	26	83.9	350	3	AAG20703	Aag20703 Arabidops
292	26	83.9	256	3	AAG09463	Arabidops	365	26	83.9	355	3	AAU57026	Aau57026 Propionib
293	26	83.9	256	3	AAW44075	Arabidops	366	26	83.9	361	6	ABM53545	Abm53545 Propionib
294	26	83.9	256	3	AAG48293	Arabidops	367	26	83.9	364	2	AAZ31014	Aaz31014 Prod. of
295	26	83.9	256	6	ABM15888	Mycobacte	368	26	83.9	364	3	AAG27535	Aag27535 Arabidops
296	26	83.9	257	5	ABG96330	Human ova	369	26	83.9	364	3	AAG27535	Aag27535 Arabidops
297	26	83.9	257	6	ABU62601	Human fol	370	26	83.9	371	3	AAG27534	Aag27534 Arabidops
298	26	83.9	258	4	AAW93113	C glutami	371	26	83.9	372	4	AAU87492	Aau87492 Novel cen
299	26	83.9	259	5	ABU50724	Helicobac	372	26	83.9	374	4	AAE04390	Aae04390 Turkey P2
300	26	83.9	261	7	ADC77517	Zebrafish	373	26	83.9	376	2	AAW46287	Aaw46287 Rhizobium
301	26	83.9	266	2	AAW21775	Protein e	374	26	83.9	377	2	AAW46288	Aaw46288 Brucella
302	26	83.9	266	2	AAW21774	Protein e	375	26	83.9	377	6	ABJ25435	Abj25435 Aspergill
303	26	83.9	266	3	AAG40842	Zea mays	376	26	83.9	377	6	ABJ25435	Abj25435 Aspergill
304	26	83.9	267	6	ADA34903	Acinetoba	377	26	83.9	380	6	ADA33210	Ada33210 Acinetoba
305	26	83.9	267	6	ABM68044	Phototrab	378	26	83.9	381	4	AAW75209	Aaw75209 Drosophil
306	26	83.9	270	5	ABP41366	Human ova	379	26	83.9	381	5	AAE30511	Aae30511 Fruit fly
307	26	83.9	271	3	AAG05658	Arabidops	380	26	83.9	383	4	AAW82666	Aaw82666 S. epider
308	26	83.9	271	3	AAG52325	Arabidops	381	26	83.9	383	6	ABJ19176	Abj19176 Pathogen
309	26	83.9	277	2	AAW85849	S. pneumo	382	26	83.9	384	5	ABP38592	Abp38592 Staphyloc
310	26	83.9	281	3	AAG52324	Arabidops	383	26	83.9	395	4	ABG04577	Abg04577 Novel hum
311	26	83.9	281	3	AAG05657	Arabidops	384	26	83.9	396	5	ABB48280	Abb48280 Listeria
312	26	83.9	281	6	ABU11755	Human MFD	385	26	83.9	400	4	AAW90734	Aaw90734 C glutami
313	26	83.9	286	2	AAW98520	H. pylori	386	26	83.9	406	3	AAW94972	Aaw94972 Human mem
314	26	83.9	286	4	ABE65174	Drosophil	387	26	83.9	406	4	AAW74711	Aaw74711 Human mem
315	26	83.9	288	7	ADE08576	Novel pro	388	26	83.9	406	4	AAW88600	Aaw88600 Human hyd
316	26	83.9	291	3	AAG11229	Arabidops	389	26	83.9	406	4	ABU52821	Abu52821 Human tra
317	26	83.9	295	3	AAG05656	Arabidops	390	26	83.9	406	5	ABB90321	Abb90321 Human pol

391	26	83.9	406	6	ABR47440	Abu47440 Breast ca	464	26	83.9	742	6	ABU04620	Human exp
392	26	83.9	406	6	ABR47441	Breast ca	465	26	83.9	742	6	ABU04645	Human exp
393	26	83.9	406	7	ADC37401	Nuclear f	466	26	83.9	742	6	ABU04616	Human exp
394	26	83.9	406	7	ADC37403	Nuclear f	467	26	83.9	750	6	ABU36066	Protein e
395	26	83.9	412	6	ABG22921	Protein e	468	26	83.9	757	2	AAU30342	A calpain
396	26	83.9	413	6	ABG74936	A. gossyp	469	26	83.9	761	2	AAU48036	Mycobacte
397	26	83.9	414	5	ABBS4432	Lactococc	470	26	83.9	766	6	ABU36065	Protein e
398	26	83.9	423	3	AAG50315	Arabidops	471	26	83.9	767	5	AAE19164	Human mem
399	26	83.9	425	3	ABAB3667	Human can	472	26	83.9	771	5	AAO14910	Human mem
400	26	83.9	426	4	AAU35705	Helicobac	473	26	83.9	773	4	AAU45423	Propionib
401	26	83.9	426	4	AAU35883	Helicobac	474	26	83.9	773	6	ABU52620	Human NOV
402	26	83.9	426	6	ABU30711	Protein e	475	26	83.9	773	6	ABM41942	Propionib
403	26	83.9	426	7	ADE76709	Helicobac	476	26	83.9	775	4	ABM71228	Propionib
404	26	83.9	426	7	ADE76711	Helicobac	477	26	83.9	784	4	ABU41839	Protein e
405	26	83.9	449	4	ABG27415	Novel hum	478	26	83.9	799	4	AAG90377	C glutam
406	26	83.9	452	3	AAG44023	Arabidops	479	26	83.9	802	6	ABU17714	Protein e
407	26	83.9	460	5	ABB92909	Herbicida	480	26	83.9	803	4	AAU34101	Staphyloc
408	26	83.9	468	3	AAG21037	Arabidops	481	26	83.9	803	5	ABB49376	Listeria
409	26	83.9	474	4	ABG21183	Novel hum	482	26	83.9	803	6	ABU32547	Protein e
410	26	83.9	475	5	ABP30445	Streptoco	483	26	83.9	804	2	AAW22709	Leucyl-CR
411	26	83.9	475	5	ABP70086	Human NOV	484	26	83.9	804	4	AAU33486	Enterococ
412	26	83.9	476	4	ABG08740	Novel hum	485	26	83.9	804	4	AAE02574	Enterococ
413	26	83.9	477	6	ABU01544	S. pneumo	486	26	83.9	804	6	ABU15896	Protein e
414	26	83.9	479	3	AAG21036	Arabidops	487	26	83.9	804	6	ABU43506	Protein e
415	26	83.9	479	5	ABP27772	Streptoco	488	26	83.9	804	6	ABU29737	Protein e
416	26	83.9	479	6	ABU00840	S. pneumo	489	26	83.9	804	6	ABU43266	Protein e
417	26	83.9	479	6	ABU45899	Protein e	490	26	83.9	804	6	ABU14575	Protein e
418	26	83.9	487	4	AAU78923	Human pro	491	26	83.9	804	6	AAE36398	Staphyloc
419	26	83.9	487	4	AAU09938	Human pro	492	26	83.9	805	4	AAU36693	Staphyloc
420	26	83.9	487	4	AAU73515	Human tra	493	26	83.9	805	6	ADA89537	Staphyloc
421	26	83.9	487	7	ABW00577	Human 326	494	26	83.9	805	6	ABM71188	Staphyloc
422	26	83.9	488	4	ABB69695	Drosophil	495	26	83.9	806	6	ABU39600	Protein e
423	26	83.9	489	3	AAG21035	Arabidops	496	26	83.9	806	6	ABU48902	Protein e
424	26	83.9	491	7	ADB65089	Human pro	497	26	83.9	811	4	ABU36282	Pseudomon
425	26	83.9	493	3	AAG50314	Arabidops	498	26	83.9	811	6	ABU38475	Protein e
426	26	83.9	495	6	ABU50602	Protein e	499	26	83.9	812	6	ABU23742	Protein e
427	26	83.9	501	6	ABU27742	Protein e	500	26	83.9	813	6	ABU24411	Protein e
428	26	83.9	516	6	ABU49230	Protein e							
429	26	83.9	520	3	AAG44022	Arabidops							
430	26	83.9	520	4	ABG27320	Novel hum							
431	26	83.9	522	3	AAG44021	Arabidops							
432	26	83.9	522	6	ABJ25385	Aspergill							
433	26	83.9	522	6	ABJ25985	Aspergill							
434	26	83.9	528	6	ADA55681	Human pro							
435	26	83.9	528	4	AAU79907	Human pro							
436	26	83.9	533	7	ADB65070	Human pro							
437	26	83.9	539	5	ABB93904	Herbicida							
438	26	83.9	573	3	AAU19119	Polypepti							
439	26	83.9	574	3	AAU26390	Protein e							
440	26	83.9	578	5	ABB89854	Human pol							
441	26	83.9	594	4	ABB61362	Drosophil							
442	26	83.9	597	4	AAU37633	Streptoco							
443	26	83.9	619	5	ABP65770	Bifidobac							
444	26	83.9	623	3	AAW02284	Candida a							
445	26	83.9	627	2	AAW62553	Protein o							
446	26	83.9	627	2	AAW62552	Shiitake							
447	26	83.9	644	6	ABU18314	Protein e							
448	26	83.9	668	6	ABB66214	Drosophil							
449	26	83.9	668	6	ABU04619	Human exp							
450	26	83.9	673	4	ABB59674	Drosophil							
451	26	83.9	675	6	ABU04618	Human exp							
452	26	83.9	687	5	ABB92504	Herbicida							
453	26	83.9	691	6	ABU04621	Human exp							
454	26	83.9	697	6	ADB09442	Alloioioc							
455	26	83.9	711	6	ADB09444	Alloioioc							
456	26	83.9	712	5	ABB91435	Herbicida							
457	26	83.9	719	6	ABU04604	Human exp							
458	26	83.9	719	6	ABU04650	Human exp							
459	26	83.9	728	4	ABB11995	Human M83							
460	26	83.9	729	4	ABB64294	Drosophil							
461	26	83.9	730	4	AAU31800	Amino aci							
462	26	83.9	742	4	ABB61574	Drosophil							
463	26	83.9	742	6	ABU04653	Human exp							

ALIGNMENTS

RESULT 1

AAU30351 ID AAU30351 standard; peptide; 15 AA.

XX AAU30351;

DT 09-NOV-1999 (first entry)

XX Epitope derived from pneumococcal surface adhesion A protein.

XX Pneumococcal surface adhesion A protein; Psal; monoclonal antibody;

XX vaccine; Streptococcus pneumoniae infection.

XX Streptococcus pneumoniae.

XX WO9945121-A1.

XX 10-SEP-1999.

XX 26-FEB-1999; 99WO-US004326.

XX 02-MAR-1998; 98US-0076565P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Carlone GM, Ades EW, Sampson JS, Tharpe JA, Zeiler JL;

XX Weisterink MAJ;

XX WPI; 1999-540849/45.

XX New peptides corresponding to Streptococcus pneumoniae Psal, used for

PT treating or preventing Streptococcus pneumoniae infection in a subject.

PS Claim 6; Page 43; 58pp; English.

XX AAY30351-54 represent immunogenic peptides which are derived from a
 CC pneumococcal surface adhesion A protein (PsaA). The specification
 CC describes monoclonal antibodies which bind epitopes of the PsaA protein
 CC (e.g. present sequence). The peptides can be used in vaccines to prevent
 CC Streptococcus pneumoniae infections. The antibodies of the invention can
 CC also be used to detect S. pneumoniae in a sample or individual

XX SQ Sequence 15 AA;

Query Match 90.3%; Score 28; DB 2; Length 15;
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6

Db 7 WTAWAF 12

RESULT 2

AAE19239
 ID AAE19239 standard; peptide; 15 AA.

XX AC AAE19239;

XX DT 21-MAY-2002 (first entry)

XX DE Streptococcus pneumoniae PsaA immunogenic peptide #1.

XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
 KW pneumococcal surface adhesin protein A; PsaA; antibacterial.

XX OS Streptococcus pneumoniae.

XX PN WO20020497-A2.

XX PD 17-JAN-2002.

XX PF 10-JUL-2001; 2001WO-US021626.

XX PR 10-JUL-2000; 2000US-00613092.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;

XX DR WPI; 2002-195762/25.

XX PT New multiple antigenic peptide for immunizing against streptococcal
 PT infections, binds to monoclonal antibody obtained in response to
 PT immunizing an animal with pneumococcal surface adhesion protein A or its
 PT fragment.

PS Claim 2; Page 56; 86pp; English.

XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
 CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
 CC obtained in response to immunizing an animal with pneumococcal surface
 CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
 CC protective immunity against S. pneumoniae infection in a subject. The
 CC present sequence is Streptococcus pneumoniae PsaA immunogenic peptide

XX SQ Sequence 15 AA;

Query Match 90.3%; Score 28; DB 5; Length 15;
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6

Db 7 WTAWAF 12

RESULT 3

AAE19247
 ID AAE19247 standard; protein; 47 AA.

XX AC AAE19247;

XX DT 21-MAY-2002 (first entry)

XX DE Tri-peptide heterogenous multiple antigenic peptide (MAP).

XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
 KW pneumococcal surface adhesin protein A; PsaA; antibacterial.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Peptide 1. .15
 /note= "PsaA immunogenic peptide #1"

FT Peptide 16. .29
 /note= "PsaA immunogenic peptide #6"

FT Peptide 30. .43
 /note= "PsaA immunogenic peptide #5"

FT Modified-site 45
 /label= Nle

FT Modified-site 47
 /label= Nle

XX PN WO20020497-A2.

XX PD 17-JAN-2002.

XX PF 10-JUL-2001; 2001WO-US021626.

XX PR 10-JUL-2000; 2000US-00613092.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;

XX DR WPI; 2002-195762/25.

XX PT New multiple antigenic peptide for immunizing against streptococcal
 PT infections, binds to monoclonal antibody obtained in response to
 PT immunizing an animal with pneumococcal surface adhesion protein A or its
 PT fragment.

XX PS Example 1b; Fig 1C; 86pp; English.

XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
 CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
 CC obtained in response to immunizing an animal with pneumococcal surface
 CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
 CC protective immunity against S. pneumoniae infection in a subject. The
 CC present sequence is a tri-peptide heterogenous MAP used to provide
 CC protection against Streptococcus pneumoniae challenge

XX SQ Sequence 47 AA;

Query Match 90.3%; Score 28; DB 5; Length 47;
 Best Local Similarity 50.0%; Pred. No. 4.1e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6

Db 7 WTAWAF 12

RESULT 4

AAU49566
 ID AAU49566 standard; protein; 57 AA.

XX AAU49566;
 AC
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #10462.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
 PI
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59545.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 10761; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 57 AA;
 XX
 Query Match 90.3%; Score 28; DB 4; Length 57;
 Best Local Similarity 50.0%; Pred. No. 4.8e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 Db 4 WSSWAF 9
 RESULT 5
 ABM46085
 ID ABM46085 standard; protein; 57 AA.

XX ABM46085;
 AC
 XX 20-OCT-2003 (first entry)
 XX
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #10761.
 XX
 KW Acne vulgaris; antisborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO2003033515-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978825.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Valliave-Douglas J;
 XX
 DR WPI; 2003-381789/36.
 DR N-PSDB; ACP64474.
 XX
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 10761; 1481pp; English.
 XX
 CC The invention relates to an isolated polynucleotide (ACF64435-ACP64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising P. acnes
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridization. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 57 AA;
 XX
 Query Match 90.3%; Score 28; DB 6; Length 57;
 Best Local Similarity 50.0%; Pred. No. 4.8e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 Db 4 WSSWAF 9

```

RESULT 6
AAE19246
ID AAE19246 standard; protein; 64 AA.
XX AC AAE19246;
XX DT 21-MAY-2002 (first entry)
XX DE Bi-peptide heterogenous multiple antigenic peptide (MAP) #2.
XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX KW pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX OS Unidentified.
XX FH Key Location/Qualifiers
FT Peptide 1..15
FT Peptide /note= "PsaA immunogenic peptide #1"
FT Peptide 16..29
FT Peptide /note= "PsaA immunogenic peptide #5"
FT Peptide 30..44
FT Peptide /note= "PsaA immunogenic peptide #1"
FT Peptide 45..59
FT Modified-site /note= "PsaA immunogenic peptide #5"
FT Modified-site 60 /label= Nle
FT Modified-site 62 /label= Nle
FT Modified-site 64 /label= Nle
FT Modified-site 66 /label= Nle
XX PN WO200204497-A2.
XX PD 17-JAN-2002.
XX PF 10-JUL-2001; 2001WO-US021626.
XX PR 10-JUL-2000; 2000US-00613092.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX DR WPI; 2002-195762/25.
XX PT New multiple antigenic peptide for immunizing against streptococcal
XX PT infections, binds to monoclonal antibody obtained in response to
XX PT immunizing an animal with pneumococcal surface adhesion protein A or its
XX PT fragment.
XX PS Example 1b; Fig 1B; 86pp; English.
XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
XX CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX CC obtained in response to immunising an animal with pneumococcal surface
XX CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX CC protective immunity against S. pneumoniae infection in a subject. The
XX CC present sequence is a bi-peptide heterogenous MAP used to provide
XX CC protection against Streptococcus pneumoniae challenge
XX SQ Sequence 64 AA;
Query Match 90.3%; Score 28; DB 5; Length 64;
Best Local Similarity 50.0%; Pred. No. 5.3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWXP 6
Db | | |
7 WTAWAF 12
RESULT 7
AAE19245
ID AAE19245 standard; protein; 66 AA.
XX AC AAE19245;
XX DT 21-MAY-2002 (first entry)
XX DE Bi-peptide heterogenous multiple antigenic peptide (MAP) #1.
XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX KW pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX OS Unidentified.
XX FH Key Location/Qualifiers
FT Peptide 1..15
FT Peptide /note= "PsaA immunogenic peptide #1"
FT Peptide 16..30
FT Peptide /note= "PsaA immunogenic peptide #2"
FT Peptide 31..45
FT Peptide /note= "PsaA immunogenic peptide #1"
FT Peptide 46..60
FT Modified-site /note= "PsaA immunogenic peptide #2"
FT Modified-site 62 /label= Nle
FT Modified-site 64 /label= Nle
FT Modified-site 66 /label= Nle
FT Modified-site 66 /label= Nle
XX PN WO200204497-A2.
XX PD 17-JAN-2002.
XX PF 10-JUL-2001; 2001WO-US021626.
XX PR 10-JUL-2000; 2000US-00613092.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX DR WPI; 2002-195762/25.
XX PT New multiple antigenic peptide for immunizing against streptococcal
XX PT infections, binds to monoclonal antibody obtained in response to
XX PT immunizing an animal with pneumococcal surface adhesion protein A or its
XX PT fragment.
XX PS Example 1b; Fig 1A; 86pp; English.
XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
XX CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX CC obtained in response to immunising an animal with pneumococcal surface
XX CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX CC protective immunity against S. pneumoniae infection in a subject. The
XX CC present sequence is a bi-peptide heterogenous MAP used to provide
XX CC protection against Streptococcus pneumoniae challenge
XX SQ Sequence 66 AA;
Query Match 90.3%; Score 28; DB 5; Length 66;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWXP 6
Db | | |
7 WTAWAF 12
RESULT 8
ABB54348
ID ABB54348 standard; protein; 87 AA.

```

```

XX AC ABB54348;
XX DT 29-AUG-2003 (revised)
XX DT 16-MAY-2002 (first entry)
XX XX Lactococcus lactis protein p1225.
XX DE Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX KW Lactococcus lactis; IL1403.
XX OS Lactococcus lactis; IL1403.
XX PN FR2807446-A1.
XX PD 12-OCT-2001.
XX PF 11-APR-2000; 2000FR-00004630.
XX PR 11-APR-2000; 2000FR-00004630.
XX PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX DR WPI; 2002-043418/06.
XX XX New nucleotide sequence useful in the identification or Lactococcus
XX PT lactis and related species.
XX FS Claim 6; SEQ ID NO 1050; 2504pp; French.
XX CC The present invention is related to a Lactococcus lactis nucleotide
XX CC sequence (ABA90521) and related proteins (AB53300-AB55621). The nucleic
XX CC acid sequence is useful in the detection and/or amplification of nucleic
XX CC acid sequence, particularly to identify Lactococcus lactis or related
XX CC species. The proteins of the invention are useful for the biosynthesis or
XX CC biodegradation of a composition of interest. The invention helps research
XX CC in lactic bacteria, particularly useful in the production of yogurt and
XX CC cheese. Note: The sequence data for this patent is based on equivalent
XX CC patent WO200177334 (published 18-OCT-2001) which is available in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
XX CC standardise OS field)
XX SQ Sequence 87 AA;

Query Match 90.3%; Score 28; DB 5; Length 87;
Best Local Similarity 50.0%; Pred. No. 6.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
Db 17 WSAWAF 22

RESULT 9
AAU64053
ID AAU64053 standard; protein; 95 AA.
XX AC AAU64053;
XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #24949.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.

```

```

XX PD 01-NOV-2001.
XX XX 20-APR-2001; 2001WO-US012865.
XX PR 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX PA (CORI-) CORIXA CORP.
XX XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX XX WPI; 2002-616774/71.
XX DR N-PSDB; AAS59638.
XX PF Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX FS Example 1; SEQ ID NO 25248; 1069pp; English.
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 95 AA;

Query Match 90.3%; Score 28; DB 4; Length 95;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
Db 2 WSTWTF 7

RESULT 10
ABM60572
ID ABM60572 standard; protein; 95 AA.
XX AC ABM60572;
XX DT 20-OCT-2003 (first entry)
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #25248.
XX KW Acne vulgaris; antisephorhoid; dermatological; antibacterial;
XX KW immunostimulant; immune response; vaccine.
XX OS Propionibacterium acnes.
XX PN WO2003033515-A1.
XX PD 24-APR-2003.

```



```

XX PF 11-OCT-2002; 2002WO-US032727.
XX PF
XX PR 15-OCT-2001; 2001US-00978825.
XX PF
XX PA (CORI-) CORIXA CORP.
XX PF
XX PI Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;
XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX PI Barth B, Valliee-Douglas J;
XX PF
XX DR WPI; 2003-381789/36.
XX DR N-PSDB; ACF64567.
XX PF
XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX PT or for stimulating an immune response specific for a P. acnes protein.
XX PF
XX PS Example 1; SEQ ID NO 25248; 1481pp; English.
XX PF
XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX CC encoding a Propionibacterium acnes protein. The invention also relates to
XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX CC immunogenic fragments of P. acnes polypeptides. The invention
XX CC additionally encompasses expression vectors and host cells comprising a
XX CC polynucleotide of the invention; antibodies against polypeptides of the
XX CC invention; fusion proteins comprising a polypeptide of the invention; a
XX CC method for stimulating an immune response specific for a P. acnes
XX CC polypeptide and an isolated T cell population comprising T cells prepared
XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,
XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or
XX CC antigen-presenting cells that express the polypeptide); a method and kit
XX CC for detecting or determining the presence or absence of P. acnes in a
XX CC patient; and a method for inhibiting the development of P. acnes in a
XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations or antigen-presenting cells that express the
XX CC polypeptides are useful for diagnosing, preventing or treating acne
XX CC vulgaris, or for stimulating an immune response specific for a P. acnes
XX CC protein. The polynucleotides can also be used as probes or primers for
XX CC nucleic acid hybridisation. The vaccine composition is useful for the
XX CC stimulation of an immune response against P. acnes, or for treating acne,
XX CC and the kit is useful for performing a diagnostic assay. The present
XX CC sequence represents a polypeptide predicted to be encoded by an ORF (open
XX CC reading frame) contained within the P. acnes polynucleotides of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX PF
XX SQ Sequence 95 AA;

Query Match 90.3%; Score 28; DB 6; Length 95;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXP 6
Db 2 WSTWTF 7

RESULT 11
ABP78009
ID ABP78009 standard; protein; 574 AA.
XX AC
XX AC ABP78009;
XX DT 07-MAR-2003 (first entry)
XX DE
XX DE N. gonorrhoeae amino acid sequence SEQ ID 2548.
XX KW Antibacterial; infection; vaccine; gene therapy.
XX OS Neisseria gonorrhoeae.
XX PT

PN WO200279243-A2.
XX 10-OCT-2002.
XX 12-FEB-2002; 2002WO-IB002069.
XX 12-FEB-2001; 2001GB-00003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Masignani V, Monaci E;
XX WPI; 2003-058415/05.
XX N-PSDB; ABZ38979.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection.
XX Disclosure; Page 374; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABP76736-ABP81046 represent nucleic acid
XX molecules of the invention
XX SQ Sequence 574 AA;

Query Match 90.3%; Score 28; DB 6; Length 574;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXP 6
Db 51 WTAWAF 56

RESULT 12
AAE26751
ID AAE26751 standard; peptide; 9 AA.
XX AC
XX AC AAE26751;
XX DT 13-DEC-2002 (first entry)
XX DE
XX DE Fibrin binding loop #3.
XX KW Fibrin binding peptide; thrombosis; pulmonary embolism; atherosclerosis;
XX myocardial infarct; ischaemia; imaging; rheumatoid arthritis; vasotropic;
XX anaemia; hypoxia; tumour; diabetic retinopathy; autoimmune disorder;
XX inflammatory disorder; angiogenesis; stroke; cerebroprotective.
XX OS
XX OS Unidentified.
XX PN WO200255544-A2.
XX 18-JUL-2002.
XX 21-DEC-2001; 2001WO-US049534.
XX 23-DEC-2000; 2000US-00747403.
XX (DYAX-) DYAX CORP.
XX Wescott CR, Beltzer JP, Sato AK;
XX WPI; 2002-666875/71.
XX Novel synthetic fibrin-binding moiety, useful for detecting, imaging or
XX localizing fibrin-containing clots by magnetic resonance imaging,

```

PT radioimaging and for treating diseases involving thrombus formation e.g.
 PT stroke.
 XX

PS Claim 4; Page 55; 89pp; English.

XX The invention relates to a synthetic fibrin binding group having affinity
 CC for fibrin. The invention is useful for detecting fibrin in a mammalian
 CC subject which involves (a) detectably labelling the binding group; (b)
 CC administering to the subject the labelled polypeptide, and (c) detecting
 CC the labelled polypeptide in the subject. The invention is useful for
 CC treating a disease involving thrombus formation eg. deep-vein thrombosis,
 CC pulmonary embolism, cardiogenic thrombosis, atherosclerosis, myocardial
 CC infarct, reperfusion ischaemia or stroke. The binding moieties are useful
 CC for detection, imaging and localisation of fibrin-containing clots by
 CC magnetic resonance imaging, radioimaging and other imaging methods and
 CC are also useful in the diagnosis and treatment of coronary conditions
 CC where fibrin plays a role. The fibrin binding moieties are useful for
 CC detecting and diagnosing numerous pathophysiological in which fibrin plays
 CC a role eg. peritoneal adhesions which often occur after surgery or
 CC inflammatory and neoplastic processes and are comprised of a fibrin
 CC network, fibroblasts, macrophages and new blood vessels; rheumatoid
 CC arthritis, lupus or septic arthritis which often have bits of fibrin
 CC containing tissues called rice bodies in the synovial fluid of their
 CC joints; thrombocytopenic purpura, a type of anaemia in which deposits in
 CC arterioles causes turbulent blood flow resulting in stress and
 CC destruction of red blood cells. The fibrin specific agents can also be
 CC used to detect hypoxia or ischaemia of heart, kidney, liver, lung, brain
 CC or other organs, as well as the detection of tumours, diabetic
 CC retinopathy, early or high-risk atherosclerosis and other autoimmune and
 CC inflammatory disorders. Fibrin specific agents also could provide both
 CC direct or surrogate markers of disease models in which hypoxia and
 CC angiogenesis are expected to play a role. The invention is also useful
 CC for screening molecular libraries. The present sequence is a fibrin
 CC binding loop
 XX

SQ Sequence 9 AA;

Query Match 87.1%; Score 27; DB 5; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
 | | |
 Db 3 WESWTF 8

RESULT 13

ID AAE26733 standard; peptide; 15 AA.

XX AAE26733;
 AC AAE26733;

DT 13-DEC-2002 (first entry)

DE Fibrin binding peptide #4.

XX Fibrin binding peptide; thrombosis; pulmonary embolism; atherosclerosis;
 KW myocardial infarct; ischaemia; imaging; rheumatoid arthritis; vasotropic;
 KW anaemia; hypoxia; tumour; diabetic retinopathy; autoimmune disorder;
 KW inflammatory disorder; angiogenesis; stroke; cerebroprotective.

OS Unidentified.

XX WO200255544-A2.

XX 18-JUL-2002.

XX 21-DEC-2001; 2001WO-US049534.

XX 23-DEC-2000; 2000US-00747403.

XX (DYAX-) DYAX CORP.

XX

PI Wescott CR, Beltzer JP, Sato AK;
 XX WPI; 2002-666875/71.

XX Novel synthetic fibrin-binding moiety, useful for detecting, imaging or
 PT localizing fibrin-containing clots by magnetic resonance imaging,
 PT radioimaging and for treating diseases involving thrombus formation e.g.
 PT stroke.

XX Claim 10; Page 57; 89pp; English.

XX The invention relates to a synthetic fibrin binding group having affinity
 CC for fibrin. The invention is useful for detecting fibrin in a mammalian
 CC subject which involves (a) detectably labelling the binding group; (b)
 CC administering to the subject the labelled polypeptide, and (c) detecting
 CC the labelled polypeptide in the subject. The invention is useful for
 CC treating a disease involving thrombus formation eg. deep-vein thrombosis,
 CC pulmonary embolism, cardiogenic thrombosis, atherosclerosis, myocardial
 CC infarct, reperfusion ischaemia or stroke. The binding moieties are useful
 CC for detection, imaging and localisation of fibrin-containing clots by
 CC magnetic resonance imaging, radioimaging and other imaging methods and
 CC are also useful in the diagnosis and treatment of coronary conditions
 CC where fibrin plays a role. The fibrin binding moieties are useful for
 CC detecting and diagnosing numerous pathophysiological in which fibrin plays
 CC a role eg. peritoneal adhesions which often occur after surgery or
 CC inflammatory and neoplastic processes and are comprised of a fibrin
 CC network, fibroblasts, macrophages and new blood vessels; rheumatoid
 CC arthritis, lupus or septic arthritis which often have bits of fibrin
 CC containing tissues called rice bodies in the synovial fluid of their
 CC joints; thrombocytopenic purpura, a type of anaemia in which deposits in
 CC arterioles causes turbulent blood flow resulting in stress and
 CC destruction of red blood cells. The fibrin specific agents can also be
 CC used to detect hypoxia or ischaemia of heart, kidney, liver, lung, brain
 CC or other organs, as well as the detection of tumours, diabetic
 CC retinopathy, early or high-risk atherosclerosis and other autoimmune and
 CC inflammatory disorders. Fibrin specific agents also could provide both
 CC direct or surrogate markers of disease models in which hypoxia and
 CC angiogenesis are expected to play a role. The invention is also useful
 CC for screening molecular libraries. The present sequence is a fibrin
 CC binding peptide
 XX

SQ Sequence 15 AA;

Query Match 87.1%; Score 27; DB 5; Length 15;
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
 | | |
 Db 6 WESWTF 11

RESULT 14

ADC9282

ID ADC9282 standard; peptide; 18 AA.

XX AC ADC9282;

DT 01-JAN-2004 (first entry)

XX Cancer-related DGI-2-binder peptide - SEQ ID 115.

XX cytostatic; cancer; gene therapy; DGI-2; DGI-5; DGI-7; DGI-9; Hras;
 KW leptin; VEGF; vascular endothelial growth factor receptor; VEGF-R1;
 KW VEGF-R2; VEGF-R3; FLT1; FMS-related tyrosine kinase 1; FLK1; KDR;
 KW kinase insert domain protein receptor; EGFR; epidermal growth factor;
 KW FGFR1; fibroblast growth factor; Tie-1.

OS Unidentified.

XX WO2003035839-A2.

XX 01-MAY-2003.

XX

XX 24-OCT-2002; 2002WO-US034021.
 XX PF
 XX XX
 XX 24-OCT-2001; 2001US-0345471P.
 XX PR
 XX (DGIB-) DGI BIOTECHNOLOGIES INC.
 XX PA
 XX Pillutla RC, Briquette R, Spruyt M, Dedova O, Blume A;
 XX PI Prendergast J, Goldstein N;
 XX PS WPI; 2003-457332/43.
 XX DR
 XX Selecting target and target binder pairs for preparing a composition for
 XX PT treating cancer by mixing in a reaction vessel phage expressing
 XX PT biological targets and phage expressing target binders.
 XX PS
 XX Claim 26; SEQ ID NO 115; 172pp; English.
 XX XX
 CC The invention relates to a novel method of selecting target and target
 CC binder pairs comprising mixing in a reaction vessel phage expressing
 CC biological targets and phage expressing target binders, each having
 CC distinguishable selection markers and selecting target and target binder
 CC pairs based on the selection markers. The molecules of the invention
 CC demonstrate cytostatic activity whilst the method may be useful for
 CC selecting target and target binder pairs for preparing a composition for
 CC treating cancer. Furthermore, the method may be utilised during gene
 CC therapy procedures. The current sequence is that of the cancer-related
 CC DGI-2-binder peptide of the invention.
 XX XX
 XX Sequence 18 AA;
 XX
 Query Match 87.1%; Score 27; DB 7; Length 18;
 Best Local Similarity 50.0%; Pred. No. 2.8e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 | | | |
 Db 6 WSSWRF 11
 RESULT 15
 AAE26744
 ID AAE26744 standard; peptide; 19 AA.
 AC AAE26744;
 XX
 XX 13-DEC-2002 (first entry)
 DT
 XX
 DE Slow dissociating fibrin binding peptide, DX-328.
 XX
 KW Fibrin binding peptide; thrombosis; pulmonary embolism; atherosclerosis;
 KW myocardial infarct; ischaemia; imaging; rheumatoid arthritis; vasotropic;
 KW anaemia; hypoxia; tumour; diabetic retinopathy; autoimmune disorder;
 KW inflammatory disorder; angiogenesis; stroke; cerebroprotective.
 XX
 OS Unidentified.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT Modified-site 19 /note= "N-terminal acetyl"
 FT Modified-site 19 /note= "C-terminal amide; HYNIC labelled"
 XX
 XX WO200255544-A2.
 XX
 XX 18-JUL-2002.
 XX
 XX 21-DEC-2001; 2001WO-US049534.
 XX PF
 XX 23-DEC-2000; 2000US-00747403.
 XX PR
 XX (DYAX-) DYAX CORP.
 XX PA
 XX

PI Wescott CR, Beltzer JP, Sato AK;
 XX WPI; 2002-666875/71.
 XX
 CC Novel synthetic fibrin-binding moiety, useful for detecting, imaging or
 CC localizing fibrin-containing clots by magnetic resonance imaging,
 CC radioimaging and for treating diseases involving thrombus formation e.g.
 CC stroke.
 XX
 XX Example 4; Page 49; 89pp; English.
 XX
 CC The invention relates to a synthetic fibrin binding group having affinity
 CC for fibrin. The invention is useful for detecting fibrin in a mammalian
 CC subject which involves (a) detectably labelling the binding group; (b)
 CC administering to the subject the labelled polypeptide, and (c) detecting
 CC the labelled polypeptide in the subject. The invention is useful for
 CC treating a disease involving thrombus formation eg. deep-vein thrombosis,
 CC pulmonary embolism, cardiogenic thrombosis, atherosclerosis, myocardial
 CC infarct, reperfusion ischaemia or stroke. The binding moieties are useful
 CC for detection, imaging and localisation of fibrin-containing clots by
 CC magnetic resonance imaging, radioimaging and other imaging methods and
 CC are also useful in the diagnosis and treatment of coronary conditions
 CC where fibrin plays a role. The fibrin binding moieties are useful for
 CC detecting and diagnosing numerous pathophysiologies in which fibrin plays
 CC a role eg. peritoneal adhesions which often occur after surgery or
 CC inflammatory and neoplastic processes and are comprised of a fibrin
 CC network, fibroblasts, macrophages and new blood vessels; rheumatoid
 CC arthritis, lupus or septic arthritis which often have bits of fibrin
 CC containing tissues called rice bodies in the synovial fluid of their
 CC joints; thrombocytopenic purpura, a type of anaemia in which deposits in
 CC arterioles causes turbulent blood flow resulting in stress and
 CC destruction of red blood cells. The fibrin specific agents can also be
 CC used to detect hypoxia or ischaemia of heart, kidney, liver, lung, brain
 CC or other organs, as well as the detection of tumours, diabetic
 CC retinopathy, early or high-risk atherosclerosis and other autoimmune and
 CC inflammatory disorders. Fibrin specific agents also could provide both
 CC direct or surrogate markers of disease models in which hypoxia and
 CC angiogenesis are expected to play a role. The invention is also useful
 CC for screening molecular libraries. The present sequence is a slow
 CC dissociating fibrin binding peptide
 XX XX
 XX Sequence 19 AA;
 XX
 Query Match 87.1%; Score 27; DB 5; Length 19;
 Best Local Similarity 50.0%; Pred. No. 3e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 | | | |
 Db 6 WESWTF 11
 RESULT 16
 AAE26745
 ID AAE26745 standard; peptide; 19 AA.
 AC AAE26745;
 XX
 XX 13-DEC-2002 (first entry)
 DT
 XX
 DE Slow dissociating fibrin binding peptide, DX-303.
 XX
 KW Fibrin binding peptide; thrombosis; pulmonary embolism; atherosclerosis;
 KW myocardial infarct; ischaemia; imaging; rheumatoid arthritis; vasotropic;
 KW anaemia; hypoxia; tumour; diabetic retinopathy; autoimmune disorder;
 KW inflammatory disorder; angiogenesis; stroke; cerebroprotective.
 XX
 OS Unidentified.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT Modified-site 19 /note= "N-terminal acetyl"
 FT Modified-site 19

/note= "C-terminal amide; fluorescein labelled"

```

FT XX Fibrin binding peptide; thrombosis; pulmonary embolism; atherosclerosis;
XX PN myocardial infarct; ischaemia; imaging; rheumatoid arthritis; vasotropic;
XX KW anaemia; hypoxia; tumour; diabetic retinopathy; autoimmune disorder;
XX KW inflammatory disorder; angiogenesis; stroke; cerebroprotective.
XX OS Unidentified.
XX XX
XX FH Key
XX FT Modified-site 1
XX FT Modified-site 19 /note= "N-terminal acetyl"
XX FT Modified-site 19 /note= "C-terminal amide"
XX XX
XX PN WO200255544-A2.
XX XX
XX PD 18-JUL-2002.
XX PD
XX PF 21-DEC-2001; 2001WO-US049534.
XX PF
XX PR 23-DEC-2000; 2000US-00747403.
XX PR
XX XX (DYAX-) DYAX CORP.
XX PA
XX XX Wescott CR, Beltzer JP, Sato AK;
XX PI
XX XX WPI; 2002-666875/71.
XX XX
XX PT Novel synthetic fibrin-binding moiety, useful for detecting, imaging or
XX PT localizing fibrin-containing clots by magnetic resonance imaging,
XX PT radioimaging and for treating diseases involving thrombus formation e.g.
XX PT stroke.
XX PS
XX PS Example 4; Page 49; 89pp; English.
XX CC The invention relates to a synthetic fibrin binding group having affinity
XX CC for fibrin. The invention is useful for detecting fibrin in a mammalian
XX CC subject which involves (a) detectably labelling the binding group; (b)
XX CC administering to the subject the labelled polypeptide, and (c) detecting
XX CC the labelled polypeptide in the subject. The invention is useful for
XX CC treating a disease involving thrombus formation eg. deep-vein thrombosis,
XX CC pulmonary embolism, cardiogenic thrombosis, atherosclerosis, myocardial
XX CC infarct, reperfusion ischaemia or stroke. The binding moieties are useful
XX CC for detection, imaging and localisation of fibrin-containing clots by
XX CC magnetic resonance imaging, radioimaging and other imaging methods and
XX CC are also useful in the diagnosis and treatment of coronary conditions
XX CC where fibrin plays a role. The fibrin binding moieties are useful for
XX CC detecting and diagnosing numerous pathophysiological in which fibrin plays
XX CC a role eg. peritoneal adhesions which often occur after surgery or
XX CC inflammatory and neoplastic processes and are comprised of a fibrin
XX CC network, fibroblasts, macrophages and new blood vessels; rheumatoid
XX CC arthritis, lupus or septic arthritis which often have bits of fibrin
XX CC containing tissues called rice bodies in the synovial fluid of their
XX CC joints; thrombocytopenic purpura, a type of anaemia in which deposits in
XX CC arterioles causes turbulent blood flow resulting in stress and
XX CC destruction of red blood cells. The fibrin specific agents can also be
XX CC used to detect hypoxia or ischaemia of heart, kidney, liver, lung, brain
XX CC or other organs, as well as the detection of tumours, diabetic
XX CC retinopathy, early or high-risk atherosclerosis and other autoimmune and
XX CC inflammatory disorders. Fibrin specific agents also could provide both
XX CC direct or surrogate markers of disease models in which hypoxia and
XX CC angiogenesis are expected to play a role. The invention is also useful
XX CC for screening molecular libraries. The present sequence is a slow
XX CC dissociating fibrin binding peptide
XX SQ
XX SQ Sequence 19 AA;
XX SQ
XX SQ Query Match 87.1%; Score 27; DB 5; Length 19;
XX SQ Best Local Similarity 50.0%; Pred. No. 3e+02;
XX SQ Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX SQ
XX QY 1 WXXWXP 6
XX DB 6 WESWTF 11
XX DB
XX RESULT 17
XX ID AAE26743
XX AC AAE26743 standard; peptide; 19 AA.
XX AC AAE26743;
XX XX
XX DT 13-DEC-2002 (first entry)
XX DE Slow dissociating fibrin binding peptide, DX-287.

```

```

XX XX Fibrin binding peptide; thrombosis; pulmonary embolism; atherosclerosis;
XX KW myocardial infarct; ischaemia; imaging; rheumatoid arthritis; vasotropic;
XX KW anaemia; hypoxia; tumour; diabetic retinopathy; autoimmune disorder;
XX KW inflammatory disorder; angiogenesis; stroke; cerebroprotective.
XX OS Unidentified.
XX XX
XX FH Key
XX FT Modified-site 1
XX FT Modified-site 19 /note= "N-terminal acetyl"
XX FT Modified-site 19 /note= "C-terminal amide"
XX XX
XX PN WO200255544-A2.
XX XX
XX PD 18-JUL-2002.
XX PD
XX PF 21-DEC-2001; 2001WO-US049534.
XX PF
XX PR 23-DEC-2000; 2000US-00747403.
XX PR
XX XX (DYAX-) DYAX CORP.
XX PA
XX XX Wescott CR, Beltzer JP, Sato AK;
XX PI
XX XX WPI; 2002-666875/71.
XX XX
XX PT Novel synthetic fibrin-binding moiety, useful for detecting, imaging or
XX PT localizing fibrin-containing clots by magnetic resonance imaging,
XX PT radioimaging and for treating diseases involving thrombus formation e.g.
XX PT stroke.
XX PS
XX PS Example 4; Page 49; 89pp; English.
XX CC The invention relates to a synthetic fibrin binding group having affinity
XX CC for fibrin. The invention is useful for detecting fibrin in a mammalian
XX CC subject which involves (a) detectably labelling the binding group; (b)
XX CC administering to the subject the labelled polypeptide, and (c) detecting
XX CC the labelled polypeptide in the subject. The invention is useful for
XX CC treating a disease involving thrombus formation eg. deep-vein thrombosis,
XX CC pulmonary embolism, cardiogenic thrombosis, atherosclerosis, myocardial
XX CC infarct, reperfusion ischaemia or stroke. The binding moieties are useful
XX CC for detection, imaging and localisation of fibrin-containing clots by
XX CC magnetic resonance imaging, radioimaging and other imaging methods and
XX CC are also useful in the diagnosis and treatment of coronary conditions
XX CC where fibrin plays a role. The fibrin binding moieties are useful for
XX CC detecting and diagnosing numerous pathophysiological in which fibrin plays
XX CC a role eg. peritoneal adhesions which often occur after surgery or
XX CC inflammatory and neoplastic processes and are comprised of a fibrin
XX CC network, fibroblasts, macrophages and new blood vessels; rheumatoid
XX CC arthritis, lupus or septic arthritis which often have bits of fibrin
XX CC containing tissues called rice bodies in the synovial fluid of their
XX CC joints; thrombocytopenic purpura, a type of anaemia in which deposits in
XX CC arterioles causes turbulent blood flow resulting in stress and
XX CC destruction of red blood cells. The fibrin specific agents can also be
XX CC used to detect hypoxia or ischaemia of heart, kidney, liver, lung, brain
XX CC or other organs, as well as the detection of tumours, diabetic
XX CC retinopathy, early or high-risk atherosclerosis and other autoimmune and
XX CC inflammatory disorders. Fibrin specific agents also could provide both
XX CC direct or surrogate markers of disease models in which hypoxia and
XX CC angiogenesis are expected to play a role. The invention is also useful
XX CC for screening molecular libraries. The present sequence is a slow
XX CC dissociating fibrin binding peptide
XX SQ
XX SQ Sequence 19 AA;
XX SQ
XX SQ Query Match 87.1%; Score 27; DB 5; Length 19;
XX SQ Best Local Similarity 50.0%; Pred. No. 3e+02;
XX SQ Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX SQ
XX QY 1 WXXWXP 6
XX DB 6 WESWTF 11
XX DB
XX RESULT 17
XX ID AAE26743
XX AC AAE26743 standard; peptide; 19 AA.
XX AC AAE26743;
XX XX
XX DT 13-DEC-2002 (first entry)
XX DE Slow dissociating fibrin binding peptide, DX-287.

```

Db 6 WESWTF 11

RESULT 18

ABG28607
ID ABG28607 standard; protein; 21 AA.

XX AC ABG28607;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #28598.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Dmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS92794.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 58966; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridization probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 21 AA;

Query Match 87.1%; Score 27; DB 4; Length 21;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6

Db 13 WAWWSF 18

RESULT 19

ABG69143
ID ABG69143 standard; peptide; 23 AA.

XX AC ABG69143;

XX DT 07-OCT-2002 (first entry)

XX DE H11 ligand binding CDR3 region from antibody A6VH11a #9.

XX KW Monoclonal antibody; A6VH-11a; phage display; panning; framework region;
XX KW Heavy chain variable region; CDR; complementarity determining region;
XX KW combinatorial library; ligand binding fragment; ligand H11.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200251870-A2.

XX PD 04-JUL-2002.

XX PF 21-DEC-2001; 2001WO-CA001845.

XX PR 22-DEC-2000; 2000US-0258031P.

XX PA (GRAD/) GRAD C.

XX PA (ENTW/) ENTWISTLE J.

XX PA (TANHA/) TANHA J.

XX PA (NARA/) NARANG S.

XX PA (DANM/) DAN M.

XX PA (MACK/) MACKENZIE C R.

XX PI Entwistle J, Tanha J, Narang S, Dan M, Mackenzie CR;

XX DR WPI; 2002-575370/61.

XX PT New libraries comprising a repertoire of nucleic acid sequences encoding
XX PT a polypeptide with a VH with CDR for expressing immunoglobulin heavy
XX PT chain binding domains for panning against the target ligand.

XX PS Disclosure; Page 19; 98pp; English.

XX CC The invention relates to a library comprising a repertoire of nucleic
CC acid sequences each encoding a polypeptide with a variable heavy (VH)
CC domain with a Complementarity Determining Region (CDR) differing at least
CC in CDR and comprising at least one of the amino acids 6-A, 23-A, 82a-N,
CC 93-A or 108-Q, in comparison to the parent human monoclonal antibody A6
CC heavy chain domain. The library (e.g. a combinatorial library for phage
CC display) may comprise at least the framework region of A6 VH including
CC the CDR region containing the target amino acids above. The libraries are
CC useful in expressing a series of potential heavy chain binding domains
CC for panning against the target ligand. The amino acid residues 6, 23,
CC 82a, 93, and 108, particularly when combined, are useful in augmenting
CC the solubility properties of a parental VH ligand binding fragment,
CC preferably one like A6, which has useful solubility properties in
CC producing a library of dbb variants for panning against a target ligand.
CC The present sequence is an H11 ligand binding peptide from a library
CC based on human monoclonal antibody A6 designated A6VH-11a, identified by
XX phage ELISA (enzyme linked immunosorbent assay)

XX SQ Sequence 23 AA;

Query Match 87.1%; Score 27; DB 5; Length 23;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6

Db 16 WGSWTF 21

RESULT 20

AAB34467
 ID AAB34467 standard; peptide; 34 AA.
 AC AAB34467;
 XX
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE Human secreted peptide #24.
 XX
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 XX WO200056767-A1.
 PN
 XX 28-SEP-2000.
 PD
 XX 16-MAR-2000; 2000WO-US006828.
 PF
 XX 19-MAR-1999; 99US-0125358P.
 PR
 PR 08-DEC-1999; 99US-0169616P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM, Komatsoulis G;
 PI
 XX WPI; 2000-602216/57.
 DR
 DR N-PSDB; AAC59652.
 XX
 PT Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers.
 XX
 PS Claim 11; Page 350; 384pp; English.
 XX
 CC Sequences AAB34439-B34484 represent the amino acid sequences of 46 human
 CC secreted proteins encoded by the genes AAC59624-C59669. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from
 CC a range of human tissues disclosed in the specification. The nucleic
 CC acids, proteins, antibodies and (ant)agonists are useful in the
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
 CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 XX infections
 XX
 SQ Sequence 34 AA;
 Query Match 87.1%; Score 27; DB 3; Length 34;
 Best Local Similarity 50.0%; Pred. No. 4.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 7 WISWAF 12
 RESULT 21
 AAM20383
 ID AAM20383 standard; protein; 52 AA.
 XX
 AC AAM20383;
 XX

12-OCT-2001 (first entry)
 Peptide #6817 encoded by probe for measuring cervical gene expression.
 Probe; human; microarray; gene expression; cervical epithelial cell;
 cervical cancer.
 Homo sapiens.
 WO200157278-A2.
 09-AUG-2001.
 30-JAN-2001; 2001WO-US000670.
 04-FEB-2000; 2000US-0180312P.
 26-MAY-2000; 2000US-0207456P.
 30-JUN-2000; 2000US-00608408.
 03-AUG-2000; 2000US-00632366.
 21-SEP-2000; 2000US-0234687P.
 27-SEP-2000; 2000US-0236359P.
 04-OCT-2000; 2000GS-00024263.
 (MOLE-) MOLECULAR DYNAMICS INC.
 Penn SG, Hanzel DK, Chen W, Rank DR;
 WPI; 2001-488901/53.
 Human genome-derived single exon nucleic acid probes useful for analyzing
 gene expression in human cervical epithelial cells.
 Claim 27; SEQ ID NO 25209; 487pp; English.
 The present invention relates to human single exon nucleic acid probes
 (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
 by one such probe. The SENPs are derived from human Hela cells. The SENPs
 can be used to produce a single exon microarray, which can be used for
 measuring human gene expression in a sample derived from human cervical
 epithelial cells. By measuring gene expression, the probes are therefore
 useful in grading and/or staging of diseases of the cervix, notably
 cervical cancer. Note: The sequence data for this patent did not form
 part of the printed specification, but was obtained in electronic format
 directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 Sequence 52 AA;
 Query Match 87.1%; Score 27; DB 4; Length 52;
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWYF 6
 Db 36 WASWNF 41
 RESULT 22
 ABB41043
 ID ABB41043 standard; peptide; 52 AA.
 XX
 AC ABB41043;
 XX
 DT 04-FEB-2002 (first entry)
 DE Peptide #8549 encoded by human foetal liver single exon probe.
 Human; foetal liver; gene expression; single exon nucleic acid probe.
 Homo sapiens.
 WO200157277-A2.
 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US0000669.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human fetal liver.
 XX PS Claim 27; SEQ ID NO 33678; 639pp + Sequence Listing; English.
 XX CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 52 AA;
 Query Match 87.1%; Score 27; DB 4; Length 52;
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 36 WASWNF 41
 RESULT 23
 AAM34818
 ID AAM34818 standard; protein; 52 AA.
 AC AAM34818;
 DT 17-OCT-2001 (first entry)
 XX PE Peptide #8955 encoded by probe for measuring placental gene expression.
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX KW genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200157272-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US0000663.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human fetal liver.
 XX PS Claim 27; SEQ ID NO 33678; 639pp + Sequence Listing; English.
 XX CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 52 AA;
 Query Match 87.1%; Score 27; DB 4; Length 52;
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 36 WASWNF 41
 RESULT 24
 ABB25116
 ID ABB25116 standard; protein; 52 AA.
 AC ABB25116;
 DT 23-JAN-2002 (first entry)
 XX PE Protein #7115 encoded by probe for measuring heart cell gene expression.
 XX KW Human; gene expression; heart; microarray; vascular system;
 XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
 XX KW congenital heart disease.
 XX OS Homo sapiens.
 XX PN WO200157274-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US0000666.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.
 XX PT Single exon nucleic acid probes for analyzing gene expression in human
 XX hearts.
 XX PS Claim 15; SEQ ID NO 26886; 530pp; English.
 XX CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human placenta.
 XX PS Claim 27; SEQ ID NO 35087; 654pp; English.
 XX CC The present invention relates to single exon nucleic acid probes (SENP:
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders
 XX SQ Sequence 52 AA;
 Query Match 87.1%; Score 27; DB 4; Length 52;
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 36 WASWNF 41

RESULT 24
 ABB25116
 ID ABB25116 standard; protein; 52 AA.
 AC ABB25116;
 DT 23-JAN-2002 (first entry)
 XX PE Protein #7115 encoded by probe for measuring heart cell gene expression.
 XX KW Human; gene expression; heart; microarray; vascular system;
 XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
 XX KW congenital heart disease.
 XX OS Homo sapiens.
 XX PN WO200157274-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US0000666.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.
 XX PT Single exon nucleic acid probes for analyzing gene expression in human
 XX hearts.
 XX PS Claim 15; SEQ ID NO 26886; 530pp; English.
 XX CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 52 AA;

Query Match 87.1%; Score 27; DB 4; Length 52;
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6
 Db 36 WASWNF 41

RESULT 25
 AAM74702
 ID AAM74702 standard; protein; 52 AA.

AC AAM74702;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 35008.

XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000668.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.

XX Example 4; SEQ ID NO 35008; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention

XX Sequence 52 AA;

Query Match 87.1%; Score 27; DB 4; Length 52;
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6

Db 36 WASWNF 41

RESULT 26

AAM61900
 ID AAM61900 standard; protein; 52 AA.

XX AC AAM61900;

XX DT 05-NOV-2001 (first entry)

XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 34005.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.

XX Example 4; SEQ ID NO 34005; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention

XX Sequence 52 AA;

Query Match 87.1%; Score 27; DB 4; Length 52;
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6

Db 36 WASWNF 41

RESULT 27

ABG56486
 ID ABG56486 standard; peptide; 52 AA.

XX AC ABG56486;

XX DT 25-FEB-2003 (first entry)

XX DE Human liver peptide, SEQ ID No 35134.

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

OS WO2000157273-A2.

PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000664.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488998/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human adult liver.

XX Claim 27; SEQ ID NO 35134; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 XX measuring human gene expression in a sample derived from human adult
 XX liver, comprising one of 13109 defined nucleotide sequences given in the
 XX specification (or complements/fragments). The probe hybridises at high
 XX stringency to a nucleic acid molecule expressed in the human adult liver.
 XX (I) may be used for predicting, measuring and displaying gene expression
 XX in samples derived from human adult liver. The genes identified may be
 XX involved in genetic liver diseases such as cirrhosis,
 XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 XX associated with coronary heart disease. ABG47348-ABG59930 represent human
 XX liver single exon encoded peptides of the invention. Note: The sequence
 XX information for this patent does not appear in the printed specification
 XX but was obtained in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 52 AA;

Query Match 87.1%; Score 27; DB 4; Length 52;

Best Local Similarity 50.0%; Pred. No. 6.6e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6

Db 36 WASWNF 41

RESULT 28

ABG44513

XX ABG44513 standard; peptide; 52 AA.

XX AC ABG44513;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 34178.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
 XX chronic obstructive pulmonary disease; interstitial lung disease;
 XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
 XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.

XX Homo sapiens.

XX WO2000186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

Spatially-addressable set of single exon nucleic acid probes, used to

measure gene expression in human lung samples.

Claim 27; SEQ ID NO 34178; 634pp; English.

The invention relates to a spatially-addressable set of single exon
 nucleic acid probes for measuring gene expression in a sample derived
 from human lung comprising single exon nucleic acid probes having one of
 12614 nucleic acid sequences mentioned in the specification, or their
 complements or the 12387 open reading frames derived from the 12614
 probes. Also included are a microarray comprising the novel set of probes
 ; the novel set of probes which hybridise at high stringency to a nucleic
 acid expressed in the human lung; measuring gene expression in a sample
 derived from human lung, comprising (a) contacting the array with a
 collection of detectably labeled nucleic acids derived from human lung
 mRNA, and (b) measuring the label detectably bound to each probe of the
 array; identifying exons in a eukaryotic genome, comprising (a)
 algorithmically predicting at least one exon from genomic sequences of
 the eukaryote; and (b) detecting specific hybridisation of detectably
 labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 having a fragment identical to the predicted exon, the probe is included
 in the above mentioned microarray; assigning exons to a single gene,
 comprising (a) identifying exons from genomic sequence by the method
 above and (b) measuring the expression of each of the exons in several
 tissues and/or cell types using hybridisation to a single exon
 microarrays having a probe with the exon, where a common pattern of
 expression of the exons in the tissues and/or cell types indicates that
 the exons should be assigned to a single gene; a peptide comprising one
 of 12011 sequences, mentioned in the specification, or encoded by the
 probes/open reading frames (ORF). The probes are used for gene expression
 analysis, and for identifying exons in a gene, particularly using human
 lung derived mRNA and for the study of lung diseases such as asthma, lung
 cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 present sequence is a peptide/protein encoded by a single exon probe of
 the invention. Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic format
 directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 52 AA;

Query Match

Best Local Similarity 87.1%; Score 27; DB 5; Length 52;

Pred. No. 6.6e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | | |
Db 36 WASWNP 41

RESULT 29
AAE21233
ID AAE21233 standard; protein; 53 AA.
AC AAE21233;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human Gene 18 encoded secreted protein HNNEM45, SEQ ID NO:98.
XX
KW Human; secreted protein; immune disorder; antiallergic; antirheumatic;
KW rheumatoid arthritis; breast neoplasia; breast cancer; antiarthritic;
KW neurological disease; Alzheimer's disease; Parkinson's disease; trauma;
KW Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania;
KW antiinflammatory; ophthalmological; dermatological; immunostimulatory;
KW immunomodulatory; immunosuppressive; antibacterial; antipsoriatic;
KW gene therapy; autoimmune disease; Huntington's disease; meningitis;
KW demyelinating disease; peripheral neuropathy; congenital malformation;
KW spinal cord injury; peripheral neuropathy; ischaemia; perception;
KW multiple sclerosis; infarction; haemorrhage; schizophrenia; dementia;
KW depression; panic disorder; learning disability; ALS; feeding disorder;
KW hyperproliferative disorder; sleep pattern; cardiovascular disorder;
KW reproductive disorder; digestive system disorder; behavioural disorder.
XX
OS Homo sapiens.
XX

Key Location/Qualifiers
FH Peptide 1..31
FT /label= Signal_peptide
FT Protein 32..53
FT /label= Mature_secretured_protein
FT Misc-difference 52
FT /label= Unknown
FT /note= "Encoded by NCA"
XX
PN WO200216390-A1.
XX
PD 28-FEB-2002.
XX
PF 17-JAN-2001; 2001WO-US001435.
XX
PR 18-AUG-2000; 2000US-0226282P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Sopbet DR;
PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
PI Fiscella M, Ni J;
XX
WPI: 2002-304113/34.
DR N-PSDB; AAD33734.
DR
XX
PT An isolated nucleic acid molecule (I) comprising a polynucleotide which
PT encodes a polypeptide useful in the diagnosis and treatment of disorders
PT e.g. immune disorders.
XX
PS Claim 11; Page 504; 504pp; English.
XX

AAD33692-AAD33736 represent cDNAs corresponding to 21 human secreted
protein genes, and AAE21191-AAE21235 represent the proteins they encode.
AAE21236-AAE21280 represent human secreted protein fragments. The genes
and their corresponding secreted proteins are useful for preventing,
treating or ameliorating medical conditions, e.g., by protein or gene
therapy. Pathological conditions can be diagnosed by determining the
amount of the new protein in a sample or by determining the presence of
mutations in the new genes. Specific uses are described for each of the

CC 21 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of immune
CC or autoimmune diseases e.g. AIDS (acquired immune deficiency syndrome),
CC asthma, anaemia and rheumatoid arthritis, breast neoplasia and breast
CC cancer, neurological diseases e.g. Alzheimer's disease, Parkinson's
CC disease, Huntington's disease, Tourette syndrome, meningitis,
CC demyelinating disease, peripheral neuropathies, neoplasia, trauma,
CC congenital malformations, spinal cord injuries, toxic neuropathies
CC induced by neurotoxins, peripheral neuropathies, multiple sclerosis,
CC ischaemia and infarction, haemorrhages, schizophrenia, mania, dementia,
CC depression, panic disorder, learning disabilities, ALS, altered
CC behaviours e.g. disorders in feeding, sleep patterns, balance and
CC perception, encephalitis, disorders in cardiovascular, neural/ sensory,
CC reproductive and digestive systems, behavioural disorders and
CC hyperproliferative disorder. The present sequence represents a human
CC secreted protein of the invention
XX
SQ Sequence 53 AA;

Query Match 87.1%; Score 27; DB 5; Length 53;
Best Local Similarity 50.0%; Pred. No. 6.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | | |
Db 21 WASWNP 26

RESULT 30
ABG64892

ID ABG64892 standard; protein; 53 AA.

AC ABG64892;

DT 27-AUG-2002 (first entry)
DE Human albumin fusion protein #1567.

XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antiinflammatory; antitumor;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.

XX Homo sapiens.
OS Synthetic.
XX
PN WO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US011988.
XX
PR 12-APR-2000; 2000US-0229358P.
PR 25-APR-2000; 2000US-019384P.
PR 21-DEC-2000; 2000US-0256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
WPI: 2002-010886/01.

New fusion protein for treating disease e.g. diabetes comprises an
albumin fused to a therapeutic protein.
PS Claim 1; Page 1582; 2102pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
XX therapeutic protein X and human albumin (HA), also known as human serum
CC

CC albumin, HSA). The proteins are useful for treating a disease or disorder
 CC that may be modulated by therapeutic protein X. The albumin extends the
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo
 CC activity. The protein is useful for treating and diagnosing disorders
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
 CC disease, ulcerative colitis), immune disorders (e.g. acquired
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
 CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
 CC ABG63326-ABG65518 represent albumin fusion proteins of the invention
 XX
 SQ Sequence 53 AA;

Query Match 87.1%; Score 27; DB 5; Length 53;

Best Local Similarity 50.0%; Pred. No. 6.7e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6

Db 21 WASHWF 26

RESULT 31

ABBI17168

ID ABBI17168 standard; protein; 56 AA.

XX AC ABBI17168;

DT 23-JAN-2002 (first entry)

DE Human nervous system related polypeptide SEQ ID NO 5825.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;

XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;

XX antiparkinsonian; antischling; antianemic; antiarthritic; cancer;

XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;

XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;

XX antiparasitic; cardiac; immune disorder; cardiovascular disorder;

XX neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX OS Homo sapiens.

XX PN WO200159063-A2.

XX PD 16-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001334.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0211535P.

PR 30-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225256P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227003P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 06-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 12-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 21-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 25-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 13-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0242221P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 01-DEC-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0254097P.
 PR 05-JAN-2001; 2001US-0254678P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-541565/60.
 DR N-PSDB; ABA13494.
 DR Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system cancers
 PT and metastases.
 PS Claim 11; SEQ ID NO 5825; 1701pp + Sequence Listing; English.
 CC The invention relates to novel genes (AB11004-ABA21534) and proteins
 CC (AB114578-AB118001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC Sequence 56 AA;
 XX SQ

Query Match 87.1%; Score 27; DB 5; Length 103;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Query Match 87.1%; Score 27; DB 4; Length 56;
 Best Local Similarity 50.0%; Pred. No. 7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 Db 9 WLSWSP 14

RESULT 32
 ABB89068
 ID ABB89068 standard; protein; 103 AA.
 XX
 AC ABB89068;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1444.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 FN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 XX 18-MAY-2001; 2001WO-US016450.
 PF
 XX 19-MAY-2000; 2000US-0205515P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Birse CE, Rosen CA;
 XX
 XX WPI; 2002-122018/16.
 DR N-PSDB; ABL89477.
 DR
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.
 PS Claim 11; SEQ ID NO 1444; 2081pp + Sequence Listing; English.
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABL89040-ABL90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC Sequence 103 AA;
 XX SQ

Query Match 87.1%; Score 27; DB 5; Length 103;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 WXXWF 6
Db 84 WTSWF 89

RESULT 33
AAB66406
ID AAB66406 standard; protein; 109 AA.

XX AC AAB66406;
XX 04-APR-2001 (first entry)
DT Human Fab clone C3 VL.

XX Human; VH; heavy chain variable region; VL; light chain variable region;
KW cytotatic; antiarthritic; antipsoriatic; antidiabetic; antirheumatic;
KW cytostatic; vasotropic; gynaecological; gene therapy;
KW epithelial glycoprotein-2; EGP-2; cancer; leukaemia; haemangioma;
KW acoustic neuroma; rheumatoid arthritis; psoriasis; angiogenesis;
KW telangiectasia; diabetes; vasculogenesis; reproductive disorder; C3.

XX Homo sapiens.

OS WO200069914-A2.

PN 23-NOV-2000.

XX 18-MAY-2000; 2000WO-GB001910.

XX 18-MAY-1999; 99GB-00011569.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

PI Hoogenboom HRJM, Reurs A, Beiboer SHWB;

XX WPI; 2001-016213/02.

XX N-PSDB; AAF29513.

PT Novel human antibody that recognizes the epithelial glycoprotein-2
PT antigen useful for diagnosis, prevention and/or treatment of cancer-
PT related diseases.

XX Claim 3; Fig 2; 86pp; English.

CC The present sequence is given in a specification relating to a novel
CC human antibody that is capable of recognising the Epithelial Glycoprotein
CC -2 (EGP-2) antigen. The antibody is useful for prevention and/or
CC treatment of conditions associated with cancer. Such conditions include
CC solid tumours, blood born tumours such as leukaemia, tumour metastasis,
CC benign tumours, for e.g. haemangiomas, acoustic neuromas, rheumatoid
CC arthritis, psoriasis, ocular angiogenic diseases, diabetic retinopathy,
CC myocardial angiogenesis, telangiectasia, wound granulation, angiogenesis,
CC diabetic neovascularisation, vasculogenesis, ovulation, menstruation and
CC placentation. The antibody is useful for in vivo imaging of cancer and
CC for adjuvant treatment of cancer. Vectors expressing the antibody are
CC useful for inducing an immunological response in a mammal in vivo. The
CC vector is selected from fowl pox virus (FPV), canary pox (ALVAC),
CC entomopox virus, vaccinia virus such as modified virus ankara (MVA) and
CC NYVAC or other non-replicating viral vector systems in order to provide a
CC dual immunotherapeutic approach

XX Sequence 109 AA;

Query Match 87.1%; Score 27; DB 4; Length 109;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 WXXWF 6
Db 94 WGTWTF 99

RESULT 34

AAM88588

ID AAM88588 standard; protein; 120 AA.

XX AAM88588;

XX 07-NOV-2001 (first entry)
DT Human immune/haematopoietic antigen SEQ ID NO:16181.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytostatic; gene therapy; vaccine; metastasis.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 11-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

XX 21-DEC-2001; 2001WO-EP015398.
 XX 22-DEC-2000; 2000EP-00870318.
 PR 04-JAN-2001; 2001EP-00870002.
 PR 03-JAN-2001; 2001EP-00870003.
 XX (JANC) JANSSEN PHARM NV.
 XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
 PI WPI; 2002-667002/71.
 DR N-PSDB; ABQ76394.
 XX New isolated nucleic acid representing a synthetic BAX-gene, useful as
 PT medicament for treating, preventing and/or alleviating yeast or fungal
 PT infections or proliferative disorders, or for preventing apoptosis in
 PT certain diseases.
 XX Claim 36; Fig 1; 344pp; English.
 XX This invention describes a novel nucleic acid representing a synthetic
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
 CC resistant yeast or fungi, identifying, or obtaining and identifying
 CC Candida spp. sequences that are differentially expressed in a pathway
 CC eventually leading to programmed cell death or identifying inhibitors or
 CC inhibitor sequences of Bax-induced cell death. The products of the
 CC invention have cytostatic, fungicide; immunosuppressive, virucide and
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
 CC antisense molecules and antibodies are useful as medicaments or in
 CC preparing a medicament for treating, preventing and/or alleviating
 CC diseases associated with yeast or fungi or proliferative disorders, such
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds
 CC or polypeptides, or the genetically modified organism are useful for
 CC preparing a medicament for modifying the endogenic flora of humans and
 CC other mammals. The vaccine is useful for immunising against yeast or
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,
 CC ischaemia, diseases related with viral infections or neurodegenerations.
 CC This sequence represents a polypeptide associated with the Bax gene
 CC described in the disclosure of the invention
 XX Sequence 129 AA;
 SQ
 Query Match 87.1%; Score 27; DB 5; Length 129;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 DB 64 WTRWSF 69
 RESULT 36
 ADB68030
 ID ADB68030 standard; protein; 145 AA.
 XX ADB68030;
 AC ADB68030;
 XX 04-DEC-2003 (first entry)
 DT Tobacco FB7-4 protein.
 DE transcription regulatory sequence; filament; stamen; pathogenic attack;
 XX tobacco; plant; FB7-4.
 KW Nicotiana tabacum.
 OS US6573428-B1.
 XX 03-JUN-2003.
 PD 20-OCT-2000; 2000US-00694127.
 PF

XX 20-OCT-1999; 99US-0160720P.
 XX (UNII) UNIV ILLINOIS FOUND.
 PA Vodkin L, Stromvik M, Sundararaman V;
 XX WPI; 2003-615482/58.
 DR Chimeric gene for plant transformation vector, comprises transcription
 XX regulatory sequence operably linked to heterologous nucleic acid
 PT sequence.
 PT Disclosure; Fig 4; 44pp; English.
 XX The invention relates to a novel chimeric gene comprising a transcription
 CC regulatory sequence operably linked to a heterologous nucleic acid
 CC sequence where the transcription regulatory sequence provides for
 CC preferential expression in the filament of a stamen. The chimeric gene of
 CC the invention may be utilised in the study and/or prevention of
 CC pathogenic attacks on soybean plants. The current sequence is that of the
 CC tobacco FB7-4 protein of the invention, which is used for homology
 CC purposes.
 XX Sequence 145 AA;
 SQ
 Query Match 87.1%; Score 27; DB 7; Length 145;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 DB 107 MWTWTF 112
 RESULT 37
 ADB93116
 ID ADB93116 standard; protein; 146 AA.
 XX ADB93116;
 AC ADB93116;
 XX 21-NOV-2002 (first entry)
 DT S. cerevisiae BAX-associated protein fragment SEQ ID 190.
 DE Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
 XX vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
 KW neurodegeneration; cell death.
 XX Saccharomyces cerevisiae.
 OS WO200264766-A2.
 XX 22-AUG-2002.
 PD 21-DEC-2001; 2001WO-EP015398.
 PF 22-DEC-2000; 2000EP-00870318.
 PR 04-JAN-2001; 2001EP-00870002.
 PR 09-JAN-2001; 2001EP-00870003.
 XX (JANC) JANSSEN PHARM NV.
 PA Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
 XX WPI; 2002-667002/71.
 DR N-PSDB; ABQ76382.
 XX New isolated nucleic acid representing a synthetic BAX-gene, useful as
 PT medicament for treating, preventing and/or alleviating yeast or fungal
 PT infections or proliferative disorders, or for preventing apoptosis in
 PT certain diseases.

XX PS Claim 36; Fig 1; 344pp; English.

XX CC This invention describes a novel nucleic acid representing a synthetic

XX CC Bax gene. The Bax gene of the invention is useful for identifying Bax-

XX CC resistant yeast or fungi, identifying, or obtaining and identifying Bax-

XX CC Candida spp. sequences that are differentially expressed in a pathway

XX CC eventually leading to programmed cell death or identifying inhibitors or

XX CC inhibitor sequences of Bax-induced cell death. The products of the

XX CC invention have cytostatic, fungicidal, immunosuppressive, virucide and

XX CC vasotropic activity and can be used in vaccines or for gene therapy. The

XX CC isolated nucleic acids, polypeptides, pharmaceutical compositions,

XX CC antisense molecules and antibodies are useful as medicaments or in

XX CC preparing a medicament for treating, preventing and/or alleviating

XX CC diseases associated with yeast or fungi or proliferative disorders, such

XX CC as cancer, or for preventing apoptosis in certain diseases. The compounds

XX CC or polypeptides, or the genetically modified organism are useful for

XX CC preparing a medicament for modifying the endogenous flora of humans and

XX CC other mammals. The vaccine is useful for immunising against yeast or

XX CC fungal infections. Apoptosis-related diseases include autoimmune disease,

XX CC ischaemia, diseases related with viral infections or neurodegenerations.

XX CC This sequence represents a polypeptide associated with the Bax gene

XX CC described in the disclosure of the invention

XX SQ Sequence 146 AA;

Query Match 87.1%; Score 27; DB 5; Length 146;

Best Local Similarity 50.0%; Pred. No. 1.5e+03;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 64 WTRWSF 69

RESULT 38

AAU95372

ID AAU95372 standard; protein; 148 AA.

XX AC AAU95372;

XX AC AAU95372;

DT 02-JUL-2002 (first entry)

XX DE Human calcium transport protein 1 splice variant A ORF 3.

XX KW Human; human leukocyte antigen; HLA; immunogen; 83P2H3; CatrP2E11;

XX KW calcium transport protein; cancer; prostate cancer; cytostatic;

XX KW chromosome 7q34; chromosome 12q24.1; T cell; B cell; open reading frame;

XX KW ORF.

XX OS Homo sapiens.

XX FN WO200214361-A2.

XX PD 21-FEB-2002.

XX PF 17-AUG-2001; 2001WO-US025782.

XX PR 17-AUG-2000; 2000US-0226329P.

XX PA (AGEN-) AGENSYS INC.

XX PI Raitano AB, Challita-Eid PM, Faris M, Saffran DC, Afar DEH;

XX PI Levin E, Hubert RS, Ge W, Jakobovits A;

XX DR N-PSDB; ABK67425.

XX DR WPI; 2002-269179/31.

XX XX Monitoring 83P2H3 gene products for monitoring the presence of cancer in

XX PT a subject, comprises determining the status of 83P2H3 gene products in a

XX PT tissue sample from the subject and comparing it to a normal sample.

XX PA Example 45; Page 217; 270pp; English.

XX CC The invention relates to monitoring 83P2H3 (a calcium transport protein

XX CC whose gene is located on chromosome 7q34) gene products in a biological

XX CC sample from a patient who has or is suspected of having cancer

XX CC (especially prostate cancer), comprises: (a) determining the status of

XX CC 83P2H3 gene products expressed by cells in a tissue sample from an

XX CC individual and (b) comparing the status to the status of 83P2H3 gene

XX CC products in a normal sample. Also included are modulators of 83P2H3

XX CC function or status, generating antibodies/immune response against 83P2H3

XX CC (or related protein CatrP2E11 whose gene is located on chromosome

XX CC 12q24.1) using identified HLA (human leukocyte antigen) binding peptides

XX CC derived from the protein, delivering a cytotoxic agent to a cell

XX CC expressing 83P2H3 by conjugating the agent to an anti-83P2H3 antibody, a

XX CC recombinant protein comprising an antigen-binding region of the antibody, a

XX CC non-human transgenic animal that produces the recombinant protein, a

XX CC hybridoma that produces the recombinant protein, a single-chain

XX CC monoclonal antibody that comprises the variable domains of the heavy and

XX CC light chains of the anti-83P2H3 antibody, a vector comprising a

XX CC polynucleotide that encodes the monoclonal antibody and inducing an

XX CC immune response to a 83P2H3 protein, by providing a 83P2H3-related

XX CC epitope that comprises a T cell or B cell epitope, and contacting the

XX CC epitope with an immune system T cell or B cell, respectively. The method

XX CC is useful for monitoring 83P2H3 gene products in a biological sample for

XX CC monitoring the presence of cancer in an individual. The modulator is

XX CC useful for inhibiting the growth of cancer cells that express 83P2H3, for

XX CC treating cancer and the vector is useful for treating a patient with a

XX CC cancer that expresses 83P2H3. The immunological methods are useful for

XX CC generating an immune response against 83P2H3, and for detecting the

XX CC presence of 83P2H3-related protein or polynucleotide in a biological

XX CC sample from a patient who has or who is suspected of having cancer. The

XX CC antibody is useful in prostate cancer diagnosis, prognosis, imaging

XX CC methodologies and treatment, to detect and quantify 83P2H3 and mutant

XX CC 83P2H3-related proteins, for purifying a 83P2H3-related protein, for

XX CC isolating 83P2H3 homologues/related molecules, and for generating anti-

XX CC idiotypic antibodies that mimic the 83P2H3 protein. The present sequence

XX CC is an open reading frame encoded by a splice variant of 83P2H3. Note: All

XX CC the X residues in the present sequence represent in frame stop codons

XX SQ Sequence 148 AA;

Query Match 87.1%; Score 27; DB 5; Length 148;

Best Local Similarity 50.0%; Pred. No. 1.5e+03;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 101 WSSWGF 106

RESULT 39

ADC01647

ID ADC01647 standard; protein; 155 AA.

XX AC ADC01647;

XX AC ADC01647;

DT 04-DEC-2003 (first entry)

XX DE Enterohaemorrhagic E. coli 0157:H7-specific protein SEQ ID NO: 1695.

XX KW enterohaemorrhagic; anti-bacterial.

XX OS Escherichia coli; 0157:H7.

XX PN JP2002355074-A.

XX PD 10-DEC-2002.

XX XX 24-JAN-2002; 2002JP-00015959.

XX PR 24-JAN-2001; 2001JP-00112010.

XX PA (UYTS-) UNIV TSUKUBA.

DR WPI; 2003-451640/43.
PT Enterohemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
XX
PS Claim 3; SEQ ID NO 1695; 2067pp; Japanese.
XX
CC The invention relates to a novel enterohaemorrhagic Escherichia coli
CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC has anti-bacterial activity. The polypeptide can be used in detection
CC and/or treatment of O157:H7 infection. The nucleotide sequence of the
CC genome of Enterohaemorrhagic E coli O157:H7 was determined. The present
CC sequence represents an E. coli O157:H7-specific polypeptide of the
CC invention.
XX
XX Sequence 155 AA;
SQ

Query Match 87.1%; Score 27; DB 7; Length 155;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWKF 6
Db 96 WWSWAF 101

RESULT 40
AAU47693
ID AAU47693 standard; protein; 163 AA.
XX
AC AAU47693;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #8589.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
FN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59539.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 8888; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 163 AA;
SQ

Query Match 87.1%; Score 27; DB 4; Length 163;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWKF 6
Db 60 MTIWAF 65

RESULT 41
ABM44212
ID ABM44212 standard; protein; 163 AA.
XX
AC ABM44212;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #8888.
XX
KW Acne vulgaris; antisborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
FN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes WJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64468.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 8888; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared

via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 163 AA;

Query Match 87.1%; Score 27; DB 6; Length 163;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | | |
Db 60 WTIAF 65

RESULT 42
ADBI1318
ID ADBI1318 standard; protein; 164 AA.
AC ADBI1318;
DT 20-NOV-2003 (first entry)
XX Alloiococcus otitis antigenic protein SEQ ID NO:5742.
DE Alloiococcus otitidis; antigenic protein; immunisation;
KW Alloiococcus otitidis; Gram-positive bacterium; infection;
KW gene therapy; Gram-positive bacterium; infection;
XX Alloiococcus otitis.
OS Alloiococcus otitis.
FN WO2003048304-A2.
XX 12-JUN-2003.
XX 25-NOV-2002; 2002WO-US036123.
XX 29-NOV-2001; 2001US-0333777P.
XX 18-NOV-2002; 2002US-0426742P.
XX (AMHP) WYETH HOLDINGS CORP.
XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX WPI; 2003-505284/47.
XX N-PSDB; ADBI1321.

New Alloiococcus otitidis polynucleotides and polypeptides, useful for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials.

Claim 33; SEQ ID NO 5742; 1019pp; English.

The present invention describes an isolated polynucleotide (I) of Alloiococcus otitidis genomic DNA, which encodes an antigenic protein. Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1) an isolated polypeptide that is encoded by the polynucleotide (I); (2) an expression vector comprising the novel isolated polynucleotide (I), its complement, degenerate variant or fragment; (3) a genetically engineered

CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Alloiococcus otitidis by administering to a host the
CC otitidis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (1) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiococcus
CC otitidis. The present sequence represents an Alloiococcus otitidis
CC antigen protein from the present invention.

XX Sequence 164 AA;

Query Match 87.1%; Score 27; DB 6; Length 164;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | | |
Db 108 WTAWIF 113

RESULT 43
ADBI1316
ID ADBI1316 standard; protein; 167 AA.
AC ADBI1316;
DT 20-NOV-2003 (first entry)
XX Alloiococcus otitis antigenic protein SEQ ID NO:5744.
DE Alloiococcus otitidis; antigenic protein; immunisation;
KW Alloiococcus otitidis; Gram-positive bacterium; infection;
KW gene therapy; Gram-positive bacterium; infection;
XX Alloiococcus otitis.
OS Alloiococcus otitis.
FN WO2003048304-A2.
XX 12-JUN-2003.
XX 25-NOV-2002; 2002WO-US036123.
XX 29-NOV-2001; 2001US-0333777P.
XX 18-NOV-2002; 2002US-0426742P.
XX (AMHP) WYETH HOLDINGS CORP.
XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX WPI; 2003-505284/47.
XX N-PSDB; ADBI1319.

New Alloiococcus otitidis polynucleotides and polypeptides, useful for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials.

Claim 33; SEQ ID NO 5744; 1019pp; English.

The present invention describes an isolated polynucleotide (I) of Alloiococcus otitidis genomic DNA, which encodes an antigenic protein. Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1) an isolated polypeptide that is encoded by the polynucleotide (I); (2) an

expression vector comprising the novel isolated polynucleotide (I), its complement, degenerate variant or fragment; (3) a genetically engineered host cell, transformed, transformed or infected with the vector of (2); (4) an antibody specific for the polypeptide of (1); (5) an immunogenic composition comprising the polypeptide, its complement, biological equivalent or fragment, or the polynucleotide that is comprised in the expression vector; (6) a pharmaceutical composition comprising the polypeptide of (1) and a carrier; (7) a protein chip comprising an array of the polypeptides of (1), their biological equivalent or fragment; (8) immunising against *Alloiooccus* otitis by administering to a host the immunogenic composition; (9) detecting and/or identifying *Alloiooccus* otitis in the biological sample; (10) a kit comprising a container containing the novel polynucleotide, its degenerate variant or fragment, or the antibody of (4); and (11) producing a polypeptide by culturing the genetically engineered host cell under conditions suitable to produce the polypeptide from the culture. (I) can be used in gene therapy. The polynucleotides, polypeptides, antibodies and compositions of the present invention can be used for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials. The polynucleotides are useful for expressing and detecting *Alloiooccus* otitis. The present sequence represents an *Alloiooccus* otitis antigen protein from the present invention.

SQ Sequence 167 AA;

Query Match 87.1%; Score 27; DB 6; Length 167;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | | |
Db 111 WTAWIF 116

RESULT 44

AAV74781
ID AAY74781 standard; protein; 169 AA.

AC AAY74781;

DT 12-SEP-2003 (revised)
DT 21-MAR-2000 (first entry)

DE *Neisseria gonorrhoeae* ORF 259 protein sequence SEQ ID NO:1036.

KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.

OS *Neisseria gonorrhoeae*.

XX WO957280-A2.

PN 11-NOV-1999.

XX 30-APR-1999; 99WO-US009346.

XX 01-MAY-1998; 98US-0083758P.

PR 31-JUL-1998; 98US-0094869P.

PR 02-SEP-1998; 98US-0098994P.

PR 02-SEP-1998; 98US-0099062P.

PR 09-OCT-1998; 98US-0103749P.

PR 09-OCT-1998; 98US-0103794P.

PR 09-OCT-1998; 98US-0103796P.

PR 25-FEB-1999; 99US-0121528P.

XX (CHIR) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;

DR WPI; 2000-062150/05.
DR N-PSDB; AAZ53543.

XX Novel *Neisseria* polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.

XX Claim 2; Page 600; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to *Neisseria*
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC *Neisseria* bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
CC field)

SQ Sequence 169 AA;

Query Match 87.1%; Score 27; DB 3; Length 169;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | | |
Db 54 WLAWAF 59

RESULT 45

ABBI7902
ID ABB17902 standard; protein; 186 AA.

AC ABB17902;

DT 23-JAN-2002 (first entry)

XX Human nervous system related polypeptide SEQ ID NO 6559.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

XX WO200159063-A2.

XX 16-AUG-2001.

PD 17-JAN-2001; 2001WO-US001334.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 28-JUN-2000; 2000US-0209467P.

PR 30-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0215647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236328P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.

PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246479P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Harash SC, Ruben SM;

WPI; 2001-541565/60.
N-PSDB; ABA14228.

Nucleic acids encoding 3224 human nervous system antigen polypeptides,
useful for preventing, diagnosing and/or treating nervous system cancers
and metastases.

Claim 11; SEQ ID NO 6559; 1701pp + Sequence Listing; English.

The invention relates to novel genes (ABA11004-ABA21534) and proteins
(AB14678-AB18001) useful for preventing, treating or ameliorating
medical conditions e.g. by protein or gene therapy. The genes are
isolated from a range of human tissues disclosed in the specification.
The nucleic acids, proteins, antibodies and (ant)agonists are useful in
the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
disorders e.g. Addison's disease, allergies, autoimmune haemolytic

CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 186 AA;

Query Match 87.1%; Score 27; DB 4; Length 186;
 Best Local Similarity 50.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
 | | |
 DB 178 WANWAF 183

RESULT 46
 ABB49316
 ID ABB49316 standard; protein; 186 AA.

XX AC ABB49316;

XX DT 05-FEB-2002 (first entry)

XX DE Listeria monocytogenes protein #2020.

XX KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 XX vitamin B12; bacterial infection; disease.

XX OS Listeria monocytogenes.

XX PN WO200177335-A2.

XX PD 18-OCT-2001.

XX PF 11-APR-2001; 2001WO-FR001119.

XX PR 11-APR-2000; 2000FR-00004629.

XX PA (INSP) INST PASTEUR.

XX PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
 PI Dussurget O, Chatouani F, Nedjari H, Glaeser P, Kunst F, Cosart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Madueno E, De Fabios B, Wehland J, Kaerst U, Entian K, Hauf J;
 XX Rose M, Voss H;

XX WPI; 2002-010914/01.

XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
 XX and prevention of *Listeria* and related bacterial infections, and related
 XX polypeptides.

XX Claim 6; SEQ ID NO 2021; 192pp; French.

XX The present invention relates to the genome sequence of *Listeria*
 XX monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 XX it are useful for selecting probes and primers for detecting genes in *L.*
 XX monocytogenes and related organisms, and for studying genetic
 XX polymorphisms and other genomes. The present sequence is a protein
 XX encoded by the genome sequence of the present invention. Proteins
 XX expressed from the genome sequence are useful for raising specific
 XX antibodies, identification of *L. monocytogenes* and related organisms,
 XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 XX B12. The genome sequence and proteins encoded by it are also useful for
 XX selecting compounds that regulate gene expression and cell replication

CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by *L.*
 CC monocytogenes and related organisms. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 186 AA;

Query Match 87.1%; Score 27; DB 5; Length 186;
 Best Local Similarity 50.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
 | | |
 DB 93 WSSWLF 98

RESULT 47
 AAY74779

ID AAY74779 standard; protein; 214 AA.

XX AC AAY74779;

XX DT 21-MAR-2000 (first entry)

XX DE *Neisseria meningitidis* ORF 259 protein sequence SEQ ID NO:1032.

XX KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 XX antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 XX antibacterial; gene therapy.

XX OS *Neisseria meningitidis*.

XX PN WO9957280-A2.

XX PD 11-NOV-1999.

XX PF 30-APR-1999; 99WO-0009346.

XX PR 01-MAY-1998; 98US-0083758P.

XX PR 31-JUL-1998; 98US-0094869P.

XX PR 02-SEP-1998; 98US-0098994P.

XX PR 02-SEP-1998; 98US-0099062P.

XX PR 09-OCT-1998; 98US-0103749P.

XX PR 09-OCT-1998; 98US-0103794P.

XX PR 09-OCT-1998; 98US-0103796P.

XX PR 25-FEB-1999; 99US-0121528P.

XX (CHIR) CHIRON CORP.

XX (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 XX Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 XX Tettelin H, Venter JC;

XX WPI; 2000-062150/05.

XX N-PSDB; AAZ53541.

XX Novel *Neisseria* polypeptides predicted to be useful antigens for
 XX vaccines and diagnostics.

XX Claim 2; Page 598-599; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
 XX represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
 XX PCR primers used in the exemplification of the present invention. The
 XX polypeptides, the polynucleotides, antibodies and compositions of the
 XX invention can be used as vaccines, as diagnostic reagents, and as
 XX immunogenic compositions. The polypeptides can be used in the manufacture
 XX of medicaments for treating or preventing infection due to *Neisseria*

CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
 CC Neisseria bacteria, or to raise antibodies. They may also be used to
 CC screen for agonists or antagonists, which may themselves have use as
 CC antibacterial agents. The polynucleotides of the invention may also be
 CC used in gene therapy protocols. The polynucleotides of the invention may also be
 CC used in gene therapy protocols.

XX SQ Sequence 214 AA;

Query Match 87.1%; Score 27; DB 3; Length 214;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXXF 6
 | | | |
 Db 54 WLAWAF 59

RESULT 48

AAAY74778
 ID AAY74778 standard; protein; 216 AA.

XX AC AAY74778;

XX DT 12-SEP-2003 (revised)
 DT 21-MAR-2000 (first entry)

XX DE Neisseria gonorrhoeae ORF 259 protein sequence SEQ ID NO:1030.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.

XX OS Neisseria gonorrhoeae.

XX PN WO9957280-A2.

XX PD 11-NOV-1999.

XX PF 30-APR-1999; 99WO-US0009346.

XX PR 01-MAY-1998; 98US-0083758P.

XX PR 31-JUL-1998; 98US-0094869P.

XX PR 02-SEP-1998; 98US-0098994P.

XX PR 02-SEP-1998; 98US-0099062P.

XX PR 09-OCT-1998; 98US-0103749P.

XX PR 09-OCT-1998; 98US-0103794P.

XX PR 09-OCT-1998; 98US-0103796P.

XX PR 25-FEB-1999; 99US-0121528P.

XX PA (CHIR) CHIRON CORP.

XX PA (GENO-) INST GENOMIC RES.

XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;

XX DR WPI; 2000-062150/05.

XX DR N-PSDB; AAZ53540.

XX PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics.

XX PS Claim 2; Page 598; 1453pp; English.

XX CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of the
 CC invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the manufacture
 CC of medicaments for treating or preventing infection due to Neisserial
 CC bacteria (e.g. meningitis and septicaemia), to detect the presence of

CC Neisseria bacteria, or to raise antibodies. They may also be used to
 CC screen for agonists or antagonists, which may themselves have use as
 CC antibacterial agents. The polynucleotides of the invention may also be
 CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
 CC field)

XX SQ Sequence 216 AA;

Query Match 87.1%; Score 27; DB 3; Length 216;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXXF 6
 | | | |
 Db 54 WLAWAF 59

RESULT 49

AAAB13554
 ID AAB13554 standard; protein; 216 AA.

XX AC AAB13554;

XX DT 06-MAR-2001 (first entry)

XX DE Streptomyces globisporus C-1027 gene cluster ORF -7.

XX KW Eneidyne C-1027 biosynthesis gene cluster; apoprotein; chromophore;
 KW open reading frame -7; cancer.

XX OS Streptomyces globisporus.

XX PN WO200040596-A1.

XX PD 13-JUL-2000.

XX PF 06-JAN-2000; 2000WO-US000446.

XX PR 06-JAN-1999; 99US-0115434P.

XX PR 05-JAN-2000; 2000US-00477962.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Shen B, Liu W, Christenson SD, Standage S;

XX DR WPI; 2000-465947/40.

XX DR N-PSDB; AAA63349, AAA63348.

XX PT Isolated nucleic acid comprising a nucleic acid encoding any of C-1027
 PT open reading frames (ORFs) -7 to 42, excluding ORF 9 (cagA), useful for
 PT the production of enediyne C-1027 antitumor antibiotics.

XX PS Claim 15; Page 78; 160pp; English.

XX CC The present sequence is the protein which is encoded by open reading
 CC frame -7 of the Streptomyces globisporus enediyne C-1027 gene cluster.
 CC Eneidyne C-1027 is an antibiotic, consisting of an apoprotein and a non-
 CC peptidic chromophore, which acts by damaging DNA. The sequences within
 CC the gene cluster, and the proteins they encode, can be used in the
 CC treatment of cancer, along with antagonists of the protein

XX SQ Sequence 216 AA;

Query Match 87.1%; Score 27; DB 3; Length 216;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXXF 6
 | | | |
 Db 97 WASWKF 102

RESULT 50

ABP80786

ID ABP80786 standard; protein; 216 AA.

XX AC ABP80786;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae amino acid sequence SEQ ID 8102.

XX KW Antibacterial; infection; vaccine; gene therapy.

XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.

XX FD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB002069.

XX PR 12-FEB-2001; 2001GB-00003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Massignani V, Monaci E;

XX DR WPI; 2003-058415/05.

XX DR N-PSDB; ABZ41756.

XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a medicament for treating or preventing N. gonorrhoeae infection.

XX PS Disclosure; Page 779; 815pp; English.

XX CC The present invention relates to proteins from Neisseria gonorrhoeae. Also disclosed are the nucleic acid molecules encoding the proteins and antibodies that specifically bind to the proteins. The composition comprising the protein, nucleic acid or antibody is useful for the manufacture of a medicament for treating or preventing N. gonorrhoeae infection, this may be in the form of a vaccine or gene therapy.

XX CC Sequences given in records ABP76736-ABP81046 represent nucleic acid molecules of the invention

XX SQ Sequence 216 AA;

Query Match

Best Local Similarity 87.1%; Score 27; DB 6; Length 216;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6

Db 54 WLAWAF 59

Search completed: June 10, 2004, 10:48:17
 Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:43:44 ; Search time 11.6667 Seconds
(without alignments)
49.470 Million cell updates/sec

Title: US-09-912-414-9

Perfect score: 31

Sequence: 1 WXXWXF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	28	90.3	21	2	PIR1731
2	28	90.3	87	2	H86753
3	28	90.3	328	2	G82301
4	28	90.3	372	2	G70783
5	28	90.3	398	2	B82894
6	28	90.3	505	2	A81806
7	28	90.3	505	2	H81084
8	28	90.3	647	2	C87693
9	28	90.3	766	2	S37894
10	27	87.1	109	2	A12029
11	27	87.1	129	2	S48902
12	27	87.1	138	2	S73390
13	27	87.1	146	2	S64569
14	27	87.1	146	2	JQ0395
15	27	87.1	151	2	S74402
16	27	87.1	155	2	D91152
17	27	87.1	155	2	G85997
18	27	87.1	157	2	S73704
19	27	87.1	172	2	S73377
20	27	87.1	186	2	A81450
21	27	87.1	186	2	AG1086
22	27	87.1	214	2	T20683
23	27	87.1	217	2	S76528
24	27	87.1	221	2	C81238
25	27	87.1	221	2	C82010
26	27	87.1	227	2	JQ1997
27	27	87.1	239	2	S13978
28	27	87.1	239	2	C84505
29	27	87.1	245	1	C65048

30	27	87.1	245	2	C85916
31	27	87.1	245	2	H91071
32	27	87.1	254	2	C28771
33	27	87.1	293	2	T32229
34	27	87.1	294	2	T05165
35	27	87.1	295	2	T40583
36	27	87.1	297	2	E83792
37	27	87.1	318	2	AB0350
38	27	87.1	350	2	S60917
39	27	87.1	351	2	S63197
40	27	87.1	357	2	A93272
41	27	87.1	381	2	T33807
42	27	87.1	406	2	JC7798
43	27	87.1	425	2	A49424
44	27	87.1	431	2	T20821
45	27	87.1	441	2	G83212
46	27	87.1	512	2	AG0842
47	27	87.1	516	2	G70556
48	27	87.1	529	2	T45254
49	27	87.1	628	2	A87596
50	27	87.1	648	2	A48646
51	27	87.1	700	2	H82428
52	27	87.1	769	2	M41029
53	27	87.1	780	2	T22555
54	27	87.1	862	2	T40934
55	27	87.1	1080	2	T43025
56	27	87.1	1154	2	T48829
57	27	87.1	1744	2	T32228
58	27	87.1	1879	2	T15962
59	26	83.9	76	2	T14724
60	26	83.9	79	2	A64750
61	26	83.9	86	2	T42185
62	26	83.9	92	2	S37509
63	26	83.9	92	2	JQ1627
64	26	83.9	101	2	S73518
65	26	83.9	106	2	C82715
66	26	83.9	108	2	AG2505
67	26	83.9	116	2	B28966
68	26	83.9	120	2	C86882
69	26	83.9	126	2	H70634
70	26	83.9	128	2	D29504
71	26	83.9	133	2	S77289
72	26	83.9	147	2	E64370
73	26	83.9	154	2	D75525
74	26	83.9	161	2	D70040
75	26	83.9	162	2	D96604
76	26	83.9	165	2	E97033
77	26	83.9	171	2	C64247
78	26	83.9	175	2	G84254
79	26	83.9	182	2	T14862
80	26	83.9	188	2	B95365
81	26	83.9	188	2	T33882
82	26	83.9	189	2	G75349
83	26	83.9	196	2	S56961
84	26	83.9	200	2	F70900
85	26	83.9	205	2	A84923
86	26	83.9	208	2	F64935
87	26	83.9	208	2	H85785
88	26	83.9	208	2	H90936
89	26	83.9	209	2	A69758
90	26	83.9	213	2	IS1413
91	26	83.9	215	2	T16542
92	26	83.9	217	2	B26411
93	26	83.9	217	2	A34295
94	26	83.9	217	2	A26411
95	26	83.9	217	2	A19644
96	26	83.9	221	2	T27980
97	26	83.9	222	2	A97575
98	26	83.9	222	2	AI2795
99	26	83.9	223	2	T37032
100	26	83.9	224	2	F70610
101	26	83.9	226	1	F64450
102	26	83.9	229	2	SS0328

hypothetical prote
hypothetical prote
reaction center pr
hypothetical prote
hypothetical prote
synaptic glycoprot
hypothetical prote
probable membrane
probable membrane
hypothetical prote
inorganic phosphat
hypothetical prote
vacuole membrane p
patterning protein
hypothetical prote
hypothetical prote
multidrug resistan
probable trpE prot
probable anthranil
hypothetical prote
amine oxidase (cop
iron(III) ABC tran
integrin beta-8 ch
hypothetical prote
probable integral
Na+/K+-exchanging
related to SRBP c
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypotet
Ig kappa chain V r
hypothetical 10.7K
probable lipoprote
conserved hypotet
hypothetical prote
Ig heavy chain pre
hypothetical prote
hypothetical prote
hypothetical 14K p
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypotet
hypothetical prote
reverse transcript
probable oxidoredu
hypothetical prote
conserved hypotet
probable membrane
hypothetical prote
hypothetical prote
probable CDP-alcoh
probable cytochrom
probable cytochrom
hypothetical prote
translation initia
hypothetical prote
translation initia
translation initia
translation initia
hypothetical prote
hypothetical prote
conserved hypotet
probable integral
hypothetical prote
cytochrome-c oxida

103	26	83.9	231	2	S69004	translation initia	176	26	83.9	365	1	MXKRAH	nonstructural prot
104	26	83.9	234	2	C64367	hypothetical prote	177	26	83.9	372	2	G82181	galactose-1-phosph
105	26	83.9	234	2	B69437	hypothetical prote	178	26	83.9	379	2	S27502	hypothetical prote
106	26	83.9	241	2	A71020	hypothetical prote	179	26	83.9	380	2	AC3494	oligopeptide trans
107	26	83.9	251	2	H90568	conserved hypotet	180	26	83.9	381	1	A53224	ubiquinol-cytochro
108	26	83.9	253	2	C95328	probable LysR-fam	181	26	83.9	381	2	S12023	ubiquinol-cytochro
109	26	83.9	254	2	E97358	inner membrane pro	182	26	83.9	381	2	S69795	UTP-hexose-1-phosp
110	26	83.9	256	2	B70750	hypothetical prote	183	26	83.9	381	2	AG2539	hypothetical prote
111	26	83.9	257	2	A45753	folate-binding pro	184	26	83.9	382	2	G90751	probable transport
112	26	83.9	259	2	AD0186	probable ferric ir	185	26	83.9	382	2	E85615	membrane protein y
113	26	83.9	270	2	C75473	probable Sec-indep	186	26	83.9	382	2	AF0612	adenine-specific m
114	26	83.9	272	1	S82900	invasion-associate	187	26	83.9	382	2	AF0612	adenine DNA methyl
115	26	83.9	272	1	T34826	hypothetical prote	188	26	83.9	386	2	A97456	phosphoserine tran
116	26	83.9	274	2	AC2299	hypothetical prote	189	26	83.9	386	2	AD2674	putative SAR DNA-b
117	26	83.9	276	2	D90595	cobalt transport A	190	26	83.9	391	2	AE3295	probable permease
118	26	83.9	281	2	H97625	hypothetical prote	191	26	83.9	394	2	D90105	transport system p
119	26	83.9	286	2	C64538	cytochrome-c oxida	192	26	83.9	395	2	E97381	phosphoserine amin
120	26	83.9	286	2	F82694	ABC transporter su	193	26	83.9	396	2	AF1354	P-loop protein - y
121	26	83.9	287	2	D81295	cytochrome-c oxida	194	26	83.9	396	2	AG1724	acyl-CoA dehydroge
122	26	83.9	291	2	AC3226	HLyD family secret	195	26	83.9	396	2	F87647	site-specific DNA-
123	26	83.9	292	2	B71970	cytochrome oxidase	196	26	83.9	402	2	T14669	probable sugar tra
124	26	83.9	297	2	F87610	hypothetical prote	197	26	83.9	403	2	B87513	probable transport
125	26	83.9	300	2	D83487	hypothetical prote	198	26	83.9	403	2	AF3432	ferredoxin reducta
126	26	83.9	300	2	G83100	hypothetical prote	199	26	83.9	404	1	B64927	redh2 protein (AJ0
127	26	83.9	301	2	AI1331	heme A farnesyltra	200	26	83.9	404	2	AE5777	argininosuccinate
128	26	83.9	301	2	AI1331	heme A farnesyltra	201	26	83.9	405	2	AE2701	chloramphenicol re
129	26	83.9	301	2	AI1331	heme A farnesyltra	202	26	83.9	405	2	AE2701	hypothetical prote
130	26	83.9	305	2	S44661	hypothetical prote	203	26	83.9	406	2	AE2701	CD44 protein - mou
131	26	83.9	312	2	H91237	hypothetical prote	204	26	83.9	406	2	AE2701	phosphopyruvate hy
132	26	83.9	312	2	D86085	hypothetical prote	205	26	83.9	406	2	AE2701	enolase - Helicoba
133	26	83.9	312	2	W86654	thymidylate syntha	206	26	83.9	407	2	AB2904	probable amino aci
134	26	83.9	314	2	AB1309	thymidylate syntha	207	26	83.9	412	2	S39734	alpha-galactosidas
135	26	83.9	314	2	AB1681	hypothetical prote	208	26	83.9	412	2	S39734	probable cell-surf
136	26	83.9	315	1	T05512	hypothetical prote	209	26	83.9	414	2	D86764	hypothetical prote
137	26	83.9	316	1	SYLBT	thymidylate syntha	210	26	83.9	414	2	S30397	phosphopyruvate hy
138	26	83.9	317	2	B86183	hypothetical prote	211	26	83.9	414	2	S30397	enolase - Helicoba
139	26	83.9	317	2	B72216	hypothetical prote	212	26	83.9	426	2	H71967	probable amino aci
140	26	83.9	317	2	H83775	endoglucanase - Th	213	26	83.9	426	2	H71967	alpha-galactosidas
141	26	83.9	318	1	YXSAT3	hypothetical prote	214	26	83.9	434	2	T47748	probable cell-surf
142	26	83.9	318	1	C89920	thymidylate syntha	215	26	83.9	436	2	G96009	hypothetical prote
143	26	83.9	320	2	G72629	thymidylate syntha	216	26	83.9	439	2	A84153	arsenical pump mem
144	26	83.9	321	2	S39586	hypothetical prote	217	26	83.9	442	2	B70069	hypothetical prote
145	26	83.9	321	2	H64877	peptide transport	218	26	83.9	447	2	T16527	cytochrome d ubiqu
146	26	83.9	321	2	F90862	peptide transport	219	26	83.9	448	2	G84146	unknown protein F2
147	26	83.9	321	2	C85756	peptide transport	220	26	83.9	451	2	E96763	conserved hypotet
148	26	83.9	321	2	AC0658	peptide transport	221	26	83.9	452	2	C69140	hypothetical prote
149	26	83.9	323	1	QJ0697	preprotein translo	222	26	83.9	455	2	G70089	NADH2 dehydrogenas
150	26	83.9	323	2	D90686	protein-export mem	223	26	83.9	460	2	S01508	argininosuccinate
151	26	83.9	323	2	H85536	protein secretion,	224	26	83.9	460	2	D97679	copper amine oxida
152	26	83.9	323	2	AH0552	protein-export mem	225	26	83.9	460	2	T48137	NADH2 dehydrogenas
153	26	83.9	324	2	G69851	conserved hypotet	226	26	83.9	460	2	T48137	phosphate ABC tran
154	26	83.9	334	2	T36052	probable cytochrom	227	26	83.9	460	2	T48137	hypothetical prote
155	26	83.9	339	2	E83419	probable permease	228	26	83.9	460	2	T48137	hypothetical prote
156	26	83.9	339	2	D75473	prolipo protein dia	229	26	83.9	471	2	T27856	glycogen synthase
157	26	83.9	341	2	A85857	probable transport	230	26	83.9	471	2	B51300	starch synthase (E
158	26	83.9	341	2	G91012	probable transport	231	26	83.9	477	2	H98000	phosphate ABC tran
159	26	83.9	341	2	B64987	hypothetical 38.1	232	26	83.9	477	2	A87285	hypothetical prote
160	26	83.9	341	2	AB0785	probable binding-p	233	26	83.9	477	2	E95055	potassium uptake p
161	26	83.9	342	2	T35135	hypothetical prote	234	26	83.9	479	2	B97925	hypothetical prote
162	26	83.9	342	2	AE0155	probable ABC trans	235	26	83.9	482	2	A43828	probable serine ca
163	26	83.9	343	2	T22018	hypothetical prote	236	26	83.9	485	2	T47469	hypothetical prote
164	26	83.9	346	2	D70558	probable cydB prot	237	26	83.9	493	2	D83814	Na+/H+ antiporter
165	26	83.9	347	2	B85432	hypothetical prote	238	26	83.9	494	2	T02523	hypothetical prote
166	26	83.9	348	1	B37760	UTP-hexose-1-phosp	239	26	83.9	495	2	AD0402	glycerol kinase [E
167	26	83.9	348	1	XNECUD	UDPglucose-hexose-	240	26	83.9	496	2	I39478	hypothetical 57.4K
168	26	83.9	348	2	B90727	galactose-1-phosph	241	26	83.9	496	2	T39478	repr protein - Str
169	26	83.9	348	2	C85578	galactose-1-phosph	242	26	83.9	496	2	S06925	replication initia
170	26	83.9	350	2	AE0594	UTP-hexose-1-phosp	243	26	83.9	498	2	T48269	nucleolar protein
171	26	83.9	352	2	A10139	probable peptide A	244	26	83.9	511	2	AI3112	hypothetical prote
172	26	83.9	352	2	D82442	hypothetical prote	245	26	83.9	511	2	D98174	oligopeptide ABC t
173	26	83.9	355	2	T23406	site-specific DNA	246	26	83.9	516	2	D82279	sodium/alanine sym
174	26	83.9	358	1	S43876	modification methy	247	26	83.9	521	2	T38466	probable NRAMP-fam
175	26	83.9	361	2	T51078	related to homeoti	248	26	83.9	522	2	D96602	nucleolar protein

249	26	83.9	524	2	T23907	hypothetical prote	322	26	83.9	1021	2	S26985	probable DNA-direc
250	26	83.9	526	2	D87532	alpha amylase fami	323	26	83.9	1096	2	C87263	hypothetical prote
251	26	83.9	528	2	T33527	hypothetical prote	324	26	83.9	1123	2	T51517	telomerase reverse
252	26	83.9	533	2	B96777	hypothetical prote	325	26	83.9	1151	2	G96805	hypothetical prote
253	26	83.9	545	2	T18694	hypothetical prote	326	26	83.9	1152	2	B84431	hypothetical prote
254	26	83.9	558	2	A42433	hypothetical prote	327	26	83.9	1185	2	A42404	probable Na+/H+ an
255	26	83.9	561	2	A26499	alpha-glucosidase	328	26	83.9	1200	2	C96025	collagen adhesin -
256	26	83.9	566	2	H84203	phosphate ABC tran	329	26	83.9	1275	2	AD0332	hypothetical prote
257	26	83.9	572	2	T50404	probable inorganic	330	26	83.9	1378	2	T47605	probable membrane
258	26	83.9	573	2	E81385	probable oxidoredu	331	26	83.9	1386	2	S73401	probable membrane
259	26	83.9	576	2	T12005	NADH2 dehydrogenas	332	26	83.9	1395	2	S68453	MG064 homolog R02
260	26	83.9	578	2	B40046	tetracycline resis	333	26	83.9	1957	2	S68453	sodium channel pro
261	26	83.9	580	1	S01830	transport system p	334	26	83.9	1998	2	T13009	hypothetical prote
262	26	83.9	580	2	T20716	hypothetical prote	335	26	83.9	2048	2	AG2109	hypothetical prote
263	26	83.9	587	2	S21139	amine oxidase (cop	336	26	83.9	2301	2	T02323	modulin-like prote
264	26	83.9	594	2	D97431	probable alpha-glu	337	26	83.9	3744	2	S46715	hypothetical prote
265	26	83.9	609	2	S65208	probable membrane	338	26	83.9	3973	2	B71612	hypothetical prote
266	26	83.9	623	2	T22177	hypothetical prote	339	25	80.6	61	2	H91196	hypothetical prote
267	26	83.9	635	2	AE2086	hypothetical prote	340	25	80.6	65	2	S23164	light-harvesting p
268	26	83.9	646	2	T41545	hypothetical trans	341	25	80.6	73	2	AH2788	hypothetical prote
269	26	83.9	650	2	G71412	probable amine oxi	342	25	80.6	83	2	A97568	hypothetical prote
270	26	83.9	666	1	S73438	MG032 homolog B01	343	25	80.6	93	2	G87098	hypothetical prote
271	26	83.9	672	1	JC7251	MG032 homolog B01	344	25	80.6	97	2	E84147	hypothetical prote
272	26	83.9	674	2	C47439	amine oxidase (cop	345	25	80.6	99	2	C84022	hypothetical prote
273	26	83.9	674	2	C47439	amine oxidase (cop	346	25	80.6	103	2	F82266	hypothetical prote
274	26	83.9	687	2	T47403	amine oxidase-like	347	25	80.6	105	2	JE0384	conserved hypotet
275	26	83.9	694	2	T33561	hypothetical prote	348	25	80.6	106	2	AG1171	NADH2 dehydrogenas
276	26	83.9	730	2	D87365	prolyl oligopeptid	349	25	80.6	108	2	AG1164	hypothetical prote
277	26	83.9	731	2	B83904	hypothetical prote	350	25	80.6	108	2	S28241	NADH2 dehydrogenas
278	26	83.9	742	2	A47195	lymphocyte homing	351	25	80.6	109	2	F83057	hypothetical prote
279	26	83.9	749	2	C84508	probable cap-bind	352	25	80.6	110	2	AD0127	conserved hypotet
280	26	83.9	750	2	C87159	cation-transportin	353	25	80.6	110	2	AG1523	hypothetical prote
281	26	83.9	750	2	S77653	cation-transportin	354	25	80.6	116	2	G82398	conserved hypotet
282	26	83.9	752	2	G70751	probable ctpB prot	355	25	80.6	117	2	A95381	conserved hypotet
283	26	83.9	752	2	E82146	copper amine oxida	356	25	80.6	118	2	S24527	hypothetical prote
284	26	83.9	756	2	T48139	probable amine oxida	357	25	80.6	119	2	S24522	Ig heavy chain v r
285	26	83.9	761	2	D70750	probable ctpA prot	358	25	80.6	119	2	S24522	Ig heavy chain v r
286	26	83.9	764	2	H83055	probable outer mem	359	25	80.6	119	2	S24513	Ig heavy chain v r
287	26	83.9	765	2	B87833	protein ZK524.3 [i	360	25	80.6	119	2	S24513	Ig heavy chain v r
288	26	83.9	766	2	F87157	probable cation-tr	361	25	80.6	119	2	S24493	Ig heavy chain v r
289	26	83.9	771	2	JC7388	M83 protein - huma	362	25	80.6	119	2	S24491	Ig heavy chain v r
290	26	83.9	780	2	S77652	cation-transportin	363	25	80.6	119	2	S24525	Ig heavy chain v r
291	26	83.9	794	1	TLEPB7	tail tubular prote	364	25	80.6	119	2	S24520	Ig heavy chain v r
292	26	83.9	802	2	T21315	hypothetical prote	365	25	80.6	119	2	S24518	Ig heavy chain v r
293	26	83.9	803	2	AD1282	leucyl-tRNA synth	366	25	80.6	119	2	S24511	Ig heavy chain v r
294	26	83.9	804	2	AH1653	leucyl-tRNA synth	367	25	80.6	119	2	S24494	Ig heavy chain v r
295	26	83.9	804	2	B89961	leucyl-tRNA synth	368	25	80.6	119	2	S24490	Ig heavy chain v r
296	26	83.9	804	2	D89650	leucyl-tRNA synth	369	25	80.6	119	2	S24515	Ig heavy chain v r
297	26	83.9	805	2	G87662	leucine-tRNA ligas	370	25	80.6	119	2	S24526	Ig heavy chain v r
298	26	83.9	806	2	A84060	peptidase, M20/M25	371	25	80.6	119	2	S24497	Ig heavy chain v r
299	26	83.9	806	2	F82899	leucyl-tRNA synth	372	25	80.6	122	2	S24517	Ig heavy chain v r
300	26	83.9	806	2	T18840	leucyl-tRNA synth	373	25	80.6	122	2	A33989	Ig heavy chain v r
301	26	83.9	811	2	B83318	hypothetical prote	374	25	80.6	128	2	S74358	Ig heavy chain v-1
302	26	83.9	812	2	D96979	cell division prot	375	25	80.6	136	2	C95132	probable membrane
303	26	83.9	821	1	CIHMH3	leucyl-tRNA synth	376	25	80.6	140	2	S32805	hypothetical prote
304	26	83.9	821	1	B34488	calpain (EC 3.4.22	377	25	80.6	144	1	AVMS67	Ig heavy chain pre
305	26	83.9	824	2	F72408	calpain (EC 3.4.22	378	25	80.6	145	2	S03844	Ig heavy chain pre
306	26	83.9	824	2	T10020	leucine-tRNA ligas	379	25	80.6	145	2	D72726	Ig heavy chain pre
307	26	83.9	829	2	H86726	leucine-tRNA ligas	380	25	80.6	147	2	AI0273	probable ribosomal
308	26	83.9	833	2	H95029	leucine-tRNA ligas	381	25	80.6	154	2	JC6036	probable membrane
309	26	83.9	833	2	C37901	leucyl-tRNA synth	382	25	80.6	154	2	E82112	integral membrane
310	26	83.9	835	2	T27887	leucine-tRNA ligas	383	25	80.6	161	2	T50092	conserved hypotet
311	26	83.9	840	2	C70131	hypothetical prote	384	25	80.6	161	1	ZPECL	hypothetical coile
312	26	83.9	868	2	T40329	leucine-tRNA ligas	385	25	80.6	164	2	F90632	signal peptidase I
313	26	83.9	869	2	S77455	hypothetical prote	386	25	80.6	164	2	B85483	prolipoprotein sig
314	26	83.9	872	2	AD2216	leucine-tRNA ligas	387	25	80.6	166	2	AB0508	signal peptidase I
315	26	83.9	878	2	C71305	leucyl-tRNA synth	388	25	80.6	166	2	S03234	lipoprotein signal
316	26	83.9	969	2	A70912	leucine-tRNA ligas	389	25	80.6	167	2	D82561	hypothetical prote
317	26	83.9	972	2	T10023	probable leus prot	390	25	80.6	169	2	A83077	probable signal pe
318	26	83.9	973	2	A85055	leucine-tRNA ligas	391	25	80.6	169	2	AI0058	probable signal pe
319	26	83.9	996	2	JE0237	probable leucyl tr	392	25	80.6	170	2	B37152	signal peptidase I
320	26	83.9	999	2	T27628	apolipoprotein E r	393	25	80.6	171	2	E82293	signal peptidase I
321	26	83.9	1008	2	T41244	hypothetical prote	394	25	80.6	173	2	E96293	NADH dehydrogenase
322	26	83.9				SECL4 protein homo		25	80.6	178	2	T29352	hypothetical prote

395 25 80.6 179 2 B83344 hypothetical prote
396 25 80.6 180 2 G82858 hypothetical prote
397 25 80.6 186 2 D82266 histidinol phosphatase
398 25 80.6 193 2 S70681 bplK protein - Bor
399 25 80.6 194 2 A82198 hypothetical prote
400 25 80.6 196 2 C90124 RNA polymerase II
401 25 80.6 199 2 H75531 cytochrome c-type
402 25 80.6 201 2 AC2147 hypothetical prote
403 25 80.6 208 2 S47687 hypothetical prote
404 25 80.6 208 2 E91158 probable enzyme [i
405 25 80.6 208 2 E86014 probable enzyme yh
406 25 80.6 208 2 AC0465 probable membrane
407 25 80.6 208 2 AC0991 probable membrane
408 25 80.6 210 2 S50984 probable membrane
409 25 80.6 217 1 H71208 hypothetical prote
410 25 80.6 219 2 A64340 conserved hypotet
411 25 80.6 222 2 G72237 conserved hypotet
412 25 80.6 227 2 G70414 conserved hypotet
413 25 80.6 230 2 D70847 hypothetical prote
414 25 80.6 231 2 A86901 hypothetical prote
415 25 80.6 234 2 D96638 hypothetical prote
416 25 80.6 235 2 B69256 conserved hypotet
417 25 80.6 239 2 T22987 conserved hypotet
418 25 80.6 242 2 AC1646 hypothetical prote
419 25 80.6 243 1 F69856 hypothetical prote
420 25 80.6 245 2 A81405 probable membrane
421 25 80.6 245 2 T21337 sec-independent pr
422 25 80.6 248 2 B10996 hypothetical prote
423 25 80.6 249 2 D71942 CDP-diacylglycerol-
424 25 80.6 251 2 T27240 hypothetical prote
425 25 80.6 253 2 B64652 hypothetical prote
426 25 80.6 254 2 E90637 conserved hypotet
427 25 80.6 254 2 E85488 hypothetical prote
428 25 80.6 254 2 A64728 hypothetical prote
429 25 80.6 255 2 AF0515 DedA family integr
430 25 80.6 256 2 G70330 hypothetical prote
431 25 80.6 257 2 G87285 hypothetical prote
432 25 80.6 257 2 A82894 conserved hypotet
433 25 80.6 260 1 G69223 molybdate-binding
434 25 80.6 260 2 S76509 hypothetical prote
435 25 80.6 262 2 S75891 hypothetical prote
436 25 80.6 262 2 S56764 hypothetical prote
437 25 80.6 262 2 AG1830 hypothetical prote
438 25 80.6 270 2 B69303 methyltransferase
439 25 80.6 276 1 BVECGG glpG protein - Esc
440 25 80.6 276 2 C91162 protein of glp reg
441 25 80.6 276 2 D86008 protein of glp reg
442 25 80.6 276 2 AE0996 probable membrane
443 25 80.6 280 2 A98805 hypothetical prote
444 25 80.6 280 2 A98805 hypothetical prote
445 25 80.6 281 2 F85613 C-5 sterol desatur
446 25 80.6 283 2 S71251 hypothetical prote
447 25 80.6 284 2 B85042 hypothetical prote
448 25 80.6 286 2 T23774 probable Chloropia
449 25 80.6 287 2 AE2765 cytochrome-c oxida
450 25 80.6 287 2 C97546 hypothetical prote
451 25 80.6 288 2 F72219 conserved hypotet
452 25 80.6 292 1 D55582 cytochrome-c oxida
453 25 80.6 296 2 D84240 hypothetical prote
454 25 80.6 297 1 S49348 cytochrome-c oxida
455 25 80.6 299 2 AH3447 cytochrome-c oxida
456 25 80.6 305 2 AG0013 cytochrome-c oxida
457 25 80.6 306 2 D70924 probable cdaA prot
458 25 80.6 308 2 T12990 probable cdaA prot
459 25 80.6 309 2 T12991 hypothetical prote
460 25 80.6 312 2 T33215 hypothetical prote
461 25 80.6 315 2 A80165 probable virulence
462 25 80.6 318 2 S75987 hypothetical prote
463 25 80.6 324 2 G82350 lipid A biosynthes
464 25 80.6 324 2 E83043 probable glycosyl
465 25 80.6 325 2 E69515 conserved hypotet
466 25 80.6 329 2 H64557 hypothetical prote
467 25 80.6 330 2 A87607 glycosyl transferase

458 25 80.6 331 2 B71949 hypothetical prote
459 25 80.6 335 2 AC2686 cation efflux syst
460 25 80.6 336 2 S61299 lipopolysaccharide
461 25 80.6 336 2 S61299 lipopolysaccharide
462 25 80.6 336 2 C81073 ADP-heptose-LPS he
463 25 80.6 338 2 H97467 probable inner mem
464 25 80.6 344 2 A86918 probable membrane
465 25 80.6 345 2 H83020 neptosyltransferas
466 25 80.6 347 2 E71254 hypothetical prote
467 25 80.6 347 2 T23773 hypothetical prote
468 25 80.6 348 2 S77598 cytochrome-c oxida
469 25 80.6 349 1 B64096 UDpglucose-hexose-
470 25 80.6 349 2 S32040 homeotic protein e
471 25 80.6 354 1 A60015 nonstructural prot
472 25 80.6 357 1 MNXRS2 nonstructural prot
473 25 80.6 357 1 JCI256 nonstructural prot
474 25 80.6 357 1 MNXRC3 nonstructural prot
475 25 80.6 359 2 I49341 MIP-1 alpha recept
476 25 80.6 364 2 T16837 hypothetical prote
477 25 80.6 366 1 XNBYUG UDpglucose-hexose-
478 25 80.6 369 2 I64234 competence locus E
479 25 80.6 370 1 XNVKUD hypothetical prote
480 25 80.6 372 2 G83574 ubiquinol-cytochro
481 25 80.6 376 1 S28664 nonstructural prot
482 25 80.6 376 2 B38891 conserved hypotet
483 25 80.6 376 2 AC0055 conserved hypotet
484 25 80.6 376 2 S45107 sugar transferase
485 25 80.6 378 2 E82839 UDpglucose-hexose-
486 25 80.6 379 1 A56685 probable membrane
487 25 80.6 382 2 AF0168 protein nhr-16 [im
488 25 80.6 387 2 G88124 nuclear receptor N
489 25 80.6 388 2 T43355

ALIGNMENTS

RESULT 1

PHI1731
Ig heavy chain V region (clone GCC-11) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Dates: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PHI1731
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PHI1675; MUID:93301607; PMID:8315385
A:Accession: PHI1731
A:Molecule type: mRNA
A:Residues: 1-21 <MCH>
A:Experimental source: B cell
A:Note: the authors translated the codon ACA for residue 13 as Ala
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match

Best Local Similarity 90.3%; Score 28; DB 2; Length 21;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6

DB 16 WTSWSP 21

RESULT 2

H86753

phage pi2 protein 25 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: H86753
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malmgren, K.; Weissbach, J.; Ehr
Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s8
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: H86753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-87 <STO>
A;Cross-references: GB:AE005176; PID:gl2723983; PIDN:AAK05130.1; GSPDB:GN00146
A;Experimental source: strain ILL403
C;Genetics:
A;Gene: pi225

Query Match 90.3%; Score 28; DB 2; Length 87;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | |
Db 17 WSAWAF 22

RESULT 3
GB2301
peptide ABC transporter, permease protein VC0619 [imported] - *Vibrio cholerae* (strain NI
A;Species: *Vibrio cholerae*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Mar-2001
A;Accession: GB2301
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
I, R.K.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: GB2301
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-328 <HEI>
A;Cross-references: GB:AE004147; GB:AE003852; NID:g9655045; PIDN:AAF933785.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain NI6961; biotype El Tor
C;Genetics:
A;Gene: VC0619
A;Map position: 1
C;Superfamily: oligopeptide permease protein oppB

Query Match 90.3%; Score 28; DB 2; Length 328;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | |
Db 180 WTSWAF 185

RESULT 4
G70783
hypothetical protein RV0906 - *Mycobacterium tuberculosis* (strain H37RV)
C;Species: *Mycobacterium tuberculosis*
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
A;Accession: G70783
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: G70783
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-372 <COL>
A;Cross-references: GB:Z73101; GB:AL123456; NID:g3261565; PIDN:CAA97381.1; PID:e241999;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: RV0906

Query Match 90.3%; Score 28; DB 2; Length 372;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | |
Db 243 WASWAF 248

RESULT 5
B82894
conserved hypothetical UUA05.1 [imported] - *Ureaplasma urealyticum*
C;Species: *Ureaplasma urealyticum*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
A;Accession: B82894
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of *Ureaplasma urealyticum*: Alternate views of a m
A;Reference number: A82870
A;Accession: B82894
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-398 <GLA>
A;Cross-references: GB:AB002138; GB:AF222894; NID:g6899390; PIDN:AAF30816.1; GSPDB:GNO
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: UUA05.1
A;Genetic code: SGC3

Query Match 90.3%; Score 28; DB 2; Length 398;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | |
Db 135 WTSWSF 140

RESULT 6
AB1806
probable two component sensor kinase NMA1803 [imported] - *Neisseria meningitidis* (stra
C;Species: *Neisseria meningitidis*
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
A;Accession: AB1806
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mor
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandrea
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A;Reference number: AB1775; MUID:20222556; PMID:10761919
A;Accession: AB1806
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-505 <PAR>
A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85029.1; PID:g7380
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1803

Query Match 90.3%; Score 28; DB 2; Length 505;
Best Local Similarity 50.0%; Pred. No. 6e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | |
Db 51 WTAWAF 55

RESULT 7
H81064
sensor histidine kinase NMB1606 [imported] - *Neisseria meningitidis* (strain MC58 serog
C;Species: *Neisseria meningitidis*
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: H81064
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: H81064
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-505 <DET>
A:Cross-references: GB:AE002511; GB:AE002098; NID:g7226857; PIDN:AAF41958.1; PID:g722685
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1606

Query Match 90.3%; Score 28; DB 2; Length 505;
Best Local Similarity 50.0%; Pred. No. 6e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 51 WTAWAF 56

RESULT 8
C87693
acetyl-CoA synthetase [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: C87693
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.F.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-647 <SFO>
A:Cross-references: GB:AE005673; NID:g13425323; PIDN:AAK25543.1; GSPDB:GN00148
C:Genetics:
A:Gene: CG3581
C:Superfamily: acetate-CoA ligase; acetate-CoA ligase homology

Query Match 90.3%; Score 28; DB 2; Length 647;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 283 WASWTF 288

RESULT 9
S37894
hypothetical protein YKL072w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein YKL352
C:Species: *Saccharomyces cerevisiae*
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 29-Oct-1999
C:Accession: S37894; S37897; S39170; S44515
R:Rasmussen, S.; von Wettstein, D.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37872
A:Accession: S37894
A:Molecule type: DNA
A:Residues: 1-766 <RAS>
A:Cross-references: EMBL:Z28072; NID:g486100; PIDN:CAA81909.1; PID:g486101; MIPS:YKL072w
A:Experimental source: strain S288C
R:Fohl, T.M.; Fohl, F.M.
submitted to the Protein Sequence Database, March 1994

A:Reference number: S37897
A:Accession: S37897
A:Molecule type: DNA
A:Residues: 1-557 <POH>
A:Cross-references: EMBL:Z28072; MIPS:YKL072w
A:Experimental source: strain S288C
R:Rasmussen, S.W.
submitted to the EMBL Data Library, November 1993
A:Reference number: S39168
A:Accession: S39170
A:Molecule type: DNA
A:Residues: 1-766 <RA2>
A:Cross-references: EMBL:X75780; NID:g433625; PIDN:CAA53402.1; PID:g433628
R:Rasmussen, S.W.
Yeast 10, 69-74, 1994
A:Title: Sequence of a 20.7 kb region of yeast chromosome XI includes the NUP100 gene,
rp in addition to seven ORFs with weak or no significant similarity to known proteins.
A:Reference number: S44513
A:Accession: S44515
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-766 <RA3>
A:Cross-references: EMBL:X75780; NID:g433625; PIDN:CAA53402.1; PID:g433628
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C:Genetics:
A:Gene: SGD:STB6
A:Cross-references: SGD:S0001555; MIPS:YKL072w
A:Map position: 11L

Query Match 90.3%; Score 28; DB 2; Length 766;
Best Local Similarity 50.0%; Pred. No. 8.3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 745 WSTWTF 750

RESULT 10
AI2029
hypothetical protein alr1791 [imported] - *Nostoc* sp. (strain PCC 7120)
C:Species: *Nostoc* sp. PCC 7120
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AI2029
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigui
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata,
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium* /
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-109 <XUR>
A:Cross-references: GB:BA000019; PIDN:BA073490.1; PID:g17130881; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1791

Query Match 87.1%; Score 27; DB 2; Length 109;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 24 WATWLF 29

RESULT 11
S48902
hypothetical protein YHR162w - Yeast (*Saccharomyces cerevisiae*)
C:Species: *Saccharomyces cerevisiae*
C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002

C;Accession: S48902
R;Macr1, C.
Submitted to the EMBL Data Library, February 1994
A;Description: The sequence of *S. cerevisiae* cosmid 9986.
A;Reference number: S46673
A;Accession: S48902
A;Molecule type: DNA
A;Residues: 1-129 <MAC>
A;Cross-references: EMBL:U00027; NID:G551319; PIDN:AAB68009.1; PID:G458905; GSPDB:GN0000
C;Genetics:
A;Gene: MIPS:YHR162w
A;Cross-references: SGD:S0001205
A;Map position: 8R
C;Superfamily: hypothetical protein YHR162w

Query Match 87.1%; Score 27; DB 2; Length 129;
Best Local Similarity 50.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | |
Db 64 WTRWSF 69

RESULT 12
S73390
hypothetical protein R02_orf138 - Mycoplasma pneumoniae (strain ATCC 29342)
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000
C;Accession: S73390
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73390
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-138 <HIM>
A;Cross-references: EMBL:AE000008; GB:U00089; NID:G1673711; PIDN:AAB95711.1; PID:G1673711
C;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: Mycoplasma pneumoniae hypothetical protein H08_orf157a

Query Match 87.1%; Score 27; DB 2; Length 138;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | |
Db 36 WSGWSF 41

RESULT 13
S64569
hypothetical protein YGR243w - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein G8620
C;Species: *Saccharomyces cerevisiae*
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
C;Accession: S64569
R;Guarriero, P.; Barreiros, T.; Azevedo, D.; Rodrigues-Pousada, C.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64565
A;Accession: S64569
A;Molecule type: DNA
A;Residues: 1-146 <GUE>
A;Cross-references: EMBL:Z73028; NID:G1323439; PIDN:CAA97272.1; PID:G1323440; GSPDB:GN00
C;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YGR243w
A;Cross-references: SGD:S0003475
A;Map position: 7R

C;Superfamily: hypothetical protein YHR162w

Query Match 87.1%; Score 27; DB 2; Length 146;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | |
Db 64 WTRWSF 69

RESULT 14
JQ0995
hypothetical 16.5K protein - common tobacco
C;Species: *Nicotiana tabacum* (common tobacco)
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 29-Sep-1999
C;Accession: JQ0995
R;Neale, A.D.; Wahleithner, J.A.; Lund, M.; Bonnett, H.T.; Kelly, A.; Meeks-Wagner, D.
Plant Cell 2, 673-684, 1990
A;Title: Chitinase, beta-1,3-glucanase, osmotin, and extensin are expressed in tobacco
A;Reference number: JQ0993; MUID:92404740; PMID:2152343
A;Accession: JQ0995
A;Molecule type: mRNA
A;Residues: 1-146 <NEA>
A;Cross-references: GB:S44872; NID:G256140; PIDN:AAB23378.1; PID:G256141
A;Experimental source: thin cell layer, cv. Samsun NN
C;Superfamily: wound-induced protein Sn-1
C;Keywords: Glycoprotein
F;41/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.1%; Score 27; DB 2; Length 146;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | |
Db 108 WMTWTF 113

RESULT 15
S74402
hypothetical protein slr0489 - *Synechocystis* sp. (strain PCC 6803)
C;Species: *Synechocystis* sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S74402
R;Kaneke, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74402
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-151 <KAN>
A;Cross-references: EMBL:D64001; GB:AB001339; NID:G1001102; PIDN:BAAL0320.1; PID:G1001
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: *Synechocystis* hypothetical protein slr0489

Query Match 87.1%; Score 27; DB 2; Length 151;
Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
| | | |
Db 44 WSWWAF 49

RESULT 16
D91152
leader peptidase HopD [imported] - *Escherichia coli* (strain O157:H7, substrain RIMD 05
C;Species: *Escherichia coli*

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C;Accession: D91152
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic
 A;Reference number: A9629; MUID:21156231; PMID:11258796
 A;Accession: D91152
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-155 <HAY>
 A;Cross-references: GB:BA000007; PIDN:BA037611.1; PID:gl13363661; GSPDB:GNO0154
 A;Experimental source: strain O157:H7, substrain RMD 0509952
 C;Genetics:
 A;Gene: ECs4188

Query Match 87.1%; Score 27; DB 2; Length 155;
 Best Local Similarity 50.0%; Pred. No. 3.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 | | | |
 Db 96 WWSWAF 101

RESULT 17

G85997
 Probable leader peptidase hopD [imported] - Escherichia coli (strain O157:H7, substrain
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C;Accession: G85997

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Pocamoussis, K.; Apodaca,
 Nature 406, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: G85997
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-155 <STO>
 A;Cross-references: GB:AE005174; NID:gl12517954; PIDN:AAG58443.1; GSPDB:GNO0145; UWGP:Z46
 A;Experimental source: strain O157:H7, substrain BD4933
 C;Genetics:
 A;Gene: hopD

Query Match 87.1%; Score 27; DB 2; Length 155;
 Best Local Similarity 50.0%; Pred. No. 3.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 | | | |
 Db 96 WWSWAF 101

RESULT 18

S73704
 Hypothetical protein H08 orf157a - Mycoplasma pneumoniae (strain ATCC 29342)
 C;Species: Mycoplasma pneumoniae
 A;Variety: ATCC 29342
 C;Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000
 C;Accession: S73704

R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A;Reference number: S73327; MUID:97105885; PMID:8948633
 A;Accession: S73704
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-157 <HIM>
 A;Cross-references: EMBL:AE000036; GB:U00089; NID:gl1674053; PIDN:AAB96026.1; PID:gl167405
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C;Genetics:
 A;Genetic code: SGC3

C;Superfamily: Mycoplasma pneumoniae hypothetical protein H08_orf157a

Query Match 87.1%; Score 27; DB 2; Length 157;
 Best Local Similarity 50.0%; Pred. No. 3.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 | | | |
 Db 37 WSGWSF 42

RESULT 19

S73377
 Hypothetical protein C09_orf172 - Mycoplasma pneumoniae (strain ATCC 29342)
 C;Species: Mycoplasma pneumoniae
 A;Variety: ATCC 29342
 C;Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 C;Accession: S73377

R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
 A;Reference number: S73327; MUID:97105885; PMID:8948633
 A;Accession: S73377

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-172 <HIM>
 A;Cross-references: EMBL:AE000006; GB:U00089; NID:gl1673695; PIDN:AAB95699.1; PID:gl1673
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C;Genetics:
 A;Genetic code: SGC3

Query Match 87.1%; Score 27; DB 2; Length 172;
 Best Local Similarity 50.0%; Pred. No. 3.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 | | | |
 Db 37 WTAWFF 42

RESULT 20

AE1450
 Hypothetical protein lin0140 [imported] - Listeria innocua (strain Clp11262)
 C;Species: Listeria innocua
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C;Accession: AE1450
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloech,
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.;
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wenlar
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AE1450

A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-186 <GLA>
 A;Cross-references: GB:AU592022; PIDN:CAC95373.1; PID:gl16412559; GSPDB:GNO0178
 A;Experimental source: strain Clp11262
 C;Genetics:
 A;Gene: lin0140

Query Match 87.1%; Score 27; DB 2; Length 186;
 Best Local Similarity 50.0%; Pred. No. 4e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 | | | |
 Db 93 WSWLFF 98

RESULT 21

AG1086

hypothetical protein lmo0094 [imported] - *Listeria monocytogenes* (strain EGD-e)
 C;Species: *Listeria monocytogenes*
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C;Accession: AG1086
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amerid, A.; Baquero, F.; Berche, P.; Bloecker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fahi, H.
 ; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A.; Title: Comparative genomics of *Listeria* species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AG1086
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-186 <GLA>
 A;Cross-references: GB:NC_003210; PIDN:CA098309.1; PID:g16409453; GSPDB:GN00177
 A;Experimental source: strain EGD-e
 C;Genetics:
 A;Gene: lmo0094

Query Match 87.1%; Score 27; DB 2; Length 186;
 Best Local Similarity 50.0%; Pred. No. 4e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 93 WSSWLF 98

RESULT 22

T20683
 hypothetical protein F10A3.7 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T20683
 R;Ilroyd, C.
 Submitted to the EMBL Data Library, March 1997
 A;Reference number: Z19309
 A;Accession: T20683
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-214 <WIL>
 A;Cross-references: EMBL:Z92829; PIDN:CA07346.1; GSPDB:GN00023; CESP:F10A3.7
 A;Experimental source: clone F10A3
 C;Genetics:
 A;Gene: CESP:F10A3.7
 A;Map position: 5
 A;Introns: 69/3; 99/1

Query Match 87.1%; Score 27; DB 2; Length 214;
 Best Local Similarity 50.0%; Pred. No. 4.4e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 6 WMSWTF 11

RESULT 23

S76528

hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
 C;Species: *Synechocystis* sp.
 A;Variety: PCC 6803
 C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C;Accession: S76528
 R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 S.
 A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S76528

A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-217 <KAN>
 A;Cross-references: EMBL:D64002; GB:AB001339; NID:g1001612; PIDN:BAA10374.1; PID:g1001
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 87.1%; Score 27; DB 2; Length 217;
 Best Local Similarity 50.0%; Pred. No. 4.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
 Db 128 WSAWIF 133

RESULT 24

C81238

hypothetical protein NMB0108 [imported] - *Neisseria meningitidis* (strain MC58 serogrou
 C;Species: *Neisseria meningitidis*
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C;Accession: C81238
 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
 A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
 A;Reference number: AB1000; MUID:20175755; PMID:10710307
 A;Accession: C81238
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-221 <TET>
 A;Cross-references: GB:AE002369; GB:AE002098; NID:g7225314; PIDN:AAF40567.1; PID:g7225
 A;Experimental source: serogroup B, strain MC58
 C;Genetics:
 A;Gene: NMB0108

Query Match 87.1%; Score 27; DB 2; Length 221;
 Best Local Similarity 50.0%; Pred. No. 4.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 54 WLAWAF 59

RESULT 25

C82010

probable integral membrane protein NMA0166 [imported] - *Neisseria meningitidis* (strain
 C;Species: *Neisseria meningitidis*
 C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C;Accession: C82010
 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mor
 Nature 404, 502-506, 2000
 A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A;Reference number: AB1775; MUID:20222556; PMID:10761919
 A;Accession: C82010
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-221 <PAR>
 A;Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83480.1; PID:g7378
 A;Experimental source: serogroup A, strain Z2491
 C;Genetics:
 A;Gene: NMA0166

Query Match 87.1%; Score 27; DB 2; Length 221;
 Best Local Similarity 50.0%; Pred. No. 4.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6


```

Db          S4 WLAWAF 59

RESULT 26
JQ1997
hypothetical 26K protein - lactic dehydrogenase virus
N:Alternate names: ORF 2 protein
C:Species: lactic dehydrogenase virus
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C:Accession: JQ1997
R:Chen, Z.; Kuo, L.; Rowland, R.R.R.; Even, C.; Faaborg, K.S.; Plagemann, P.G.W.
J. Gen. Virol. 74, 643-660, 1993
A:Title: Sequences of 3' end of genome and of 5' end of open reading frame 1a of lactate
s.
A:Reference number: JQ1990; MUID:93224885; PMID:8385693
A:Accession: JQ1997
A:Molecule type: mRNA
A:Residues: 1-227 <CHE>
A:Cross-references: GB:L06811
A:Experimental source: isolate P

Query Match      87.1%; Score 27; DB 2; Length 227;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWXXF 6
      | | | |
Db      14 WFSWAF 19

RESULT 27
S13978
artemin - brine shrimp
C:Species: Artemia sp. (brine shrimp)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C:Accession: S13978
R:de Graaf, J.; Amons, R.; Moeller, W.
Eur. J. Biochem. 193, 737-750, 1990
A:Title: The primary structure of artemin from Artemia cysts.
A:Reference number: S13978; MUID:91065380; PMID:2249691
A:Accession: S13978
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-229 <GRA>

Query Match      87.1%; Score 27; DB 2; Length 229;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWXXF 6
      | | | |
Db      50 WSNWAF 55

RESULT 28
C84505
hypothetical protein At2g13070 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84505
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VarAken, S.E.; Umayam, L.; Tallon, L.;
eaus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84505
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-239 <STO>
A:Cross-references: GB:AE002093; NID:g4586067; PIDN:AND25684.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g13070

A:Map position: 2

Query Match      87.1%; Score 27; DB 2; Length 239;
Best Local Similarity 50.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWXXF 6
      | | | |
Db      16 WALWAF 21

RESULT 29
C85048
hypothetical protein b2682 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C:Accession: C85048
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C85048
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-245 <BLAT>
A:Cross-references: GB:AB000353; GB:U00096; NID:g1789037; PIDN:AAC75729.1; PID:g178903;
A:Experimental source: strain K-12, substrain MGI655
C:Superfamily: hypothetical protein b2682

Query Match      87.1%; Score 27; DB 1; Length 245;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWXXF 6
      | | | |
Db      149 WSSWVF 154

RESULT 30
C85916
hypothetical protein Z3983 [imported] - Escherichia coli (strain O157:H7, substrain EI
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
C:Accession: C85916
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May;
iller, L.; Grotbeck, R.J.; Davis, N.W.; Lim, A.; Dialanta, E.; Potamousis, K.; Apoda;
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85916
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-245 <STO>
A:Cross-references: GB:AE005174; NID:g12517123; PIDN:AAG57791.1; GSPDB:GN00145; UMGPI;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z3983
C:Superfamily: hypothetical protein b2682

Query Match      87.1%; Score 27; DB 2; Length 245;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWXXF 6
      | | | |
Db      149 WSSWVF 154

RESULT 31
H91071
hypothetical protein ECs3544 [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli

```

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 01-Mar-2002
 C;Accession: H91071
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: H91071
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-245 <HAY>
 A;Cross-references: GB:BA000007; PIDN:BA836967.1; PID:g13363015; GSPDB:GN00154
 A;Experimental source: strain O157:H7, substrain RIMD 0509952
 C;Genetics:
 A;Gene: EC3544
 C;Superfamily: hypothetical protein b2682

Query Match 87.1%; Score 27; DB 2; Length 245;
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6
 Db 149 WSSWVF 154

RESULT 32
 C28771
 reaction center protein chain H - Rhodobacter capsulatus
 C;Species: Rhodobacter capsulatus
 C;Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 08-Oct-1999
 C;Accession: C28771
 R;Youvan, D.C.; Bylina, E.J.; Alberti, M.; Begusch, H.; Hearst, J.E.
 Cell 37, 949-957, 1984
 A;Title: Nucleotide and deduced polypeptide sequences of the photosynthetic reaction-center protein of Rhodospirillum rubrum
 A;Reference number: A90850; MUID:84259352; PMID:674416
 A;Accession: C28771
 A;Molecule type: DNA
 A;Residues: 1-254 <YOU>
 A;Cross-references: GB:K01183; NID:g152014; PIDN:AAA26171.1; PID:g455345
 C;Superfamily: photoreaction center protein H
 C;Keywords: transmembrane protein

Query Match 87.1%; Score 27; DB 2; Length 254;
 Best Local Similarity 50.0%; Pred. No. 5.1e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6
 Db 18 WSWAF 23

RESULT 33
 T23229
 hypothetical protein T23B12.10 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C;Accession: T32229
 R;Davidson, S.; Wohldmann, P.; Gillam, B.
 submitted to the EMBL Data Library, September 1997
 A;Description: The sequence of C. elegans cosmid T23B12.
 A;Reference number: Z21137
 A;Accession: T32229
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-293 <DAV>
 A;Cross-references: EMBL:AF022982; PIDN:AAB69940.1; GSPDB:GN00023; CESP:T23B12.10
 A;Experimental source: strain Bristol N2; clone T23B12
 C;Genetics:
 A;Gene: CESP:T23B12.10
 A;Map position: 5
 A;Introns: 134/1; 189/2; 214/3; 276/1

Query Match 87.1%; Score 27; DB 2; Length 293;
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6
 Db 44 WTIWTF 49

RESULT 34
 T05165
 hypothetical protein F18E5.190 - Arabidopsis thaliana
 N;Alternate names: hypothetical protein F17L22.30
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
 C;Accession: T05165; T05832
 R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; M submitted to the Protein Sequence Database, August 1998
 A;Reference number: Z15400
 A;Accession: T05165
 A;Molecule type: DNA
 A;Residues: 1-294 <BEV>
 A;Cross-references: EMBL:AL022603
 A;Experimental source: cultivar Columbia; BAC clone F18E5
 R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; submitted to the Protein Sequence Database, February 1999
 A;Reference number: Z15454
 A;Accession: T05832
 A;Molecule type: DNA
 A;Residues: 1-294 <BEW>
 A;Cross-references: EMBL:AL035527
 A;Experimental source: cultivar Columbia; BAC clone F17L22
 C;Genetics:
 A;Map position: 4
 A;Introns: 89/3; 130/3; 227/3
 A;Note: F18E5.190; F17L22.30

Query Match 87.1%; Score 27; DB 2; Length 294;
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6
 Db 174 WLSWTF 179

RESULT 35
 T40583
 synaptic glycoprotein sc2 homolog - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999
 C;Accession: T40583
 R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 1999
 A;Reference number: Z21938
 A;Accession: T40583
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-295 <SEE>
 A;Cross-references: EMBL:AL035216; PIDN:CAA22811.1; GSPDB:GN00067; SPDB:SPBC646.07C
 A;Experimental source: strain 972h; cosmid c646
 C;Genetics:
 A;Gene: SPDB:SPBC646.07C
 A;Map position: 2

Query Match 87.1%; Score 27; DB 2; Length 295;
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6
 Db 252 WASWIF 257

RESULT 36

E83792
 hypothetical protein BH1141 [imported] - Bacillus halodurans (strain C-125)
 C;Species: Bacillus halodurans
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C;Accession: E83792
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A;Reference number: A83650; MUID:20512582; PMID:11058132
 A;Accession: E83792
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-297 <STO>
 A;Cross-references: GB:AP001511; GB:BA000004; NID:gl0173727; PIDN:BA04860.1; GSPDB:GN00
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: BH1141

Query Match 87.1%; Score 27; DB 2; Length 297;
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 | | | | |
 Db 74 WAAWIF 79

RESULT 37

AB0350
 probable membrane protein YPO2874 [imported] - Yersinia pestis (strain CO92)
 C;Species: Yersinia pestis
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
 C;Accession: AB0350
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
 Nature 413, 523-527, 2001
 A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A;Reference number: AB0001; MUID:21470413; PMID:11586360
 A;Accession: AB0350
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-318 <KUR>
 A;Cross-references: GB:AL590842; PIDN:CAC92125.1; PID:gl5980841; GSPDB:GN00175
 C;Genetics:
 A;Gene: YPO2874

Query Match 87.1%; Score 27; DB 2; Length 318;
 Best Local Similarity 50.0%; Pred. No. 6.1e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 | | | | |
 Db 3 WETWSF 8

RESULT 38

S60917
 Probable membrane protein YNL264c - yeast (Saccharomyces cerevisiae)
 N;Alternate names: hypothetical protein N0815
 C;Species: Saccharomyces cerevisiae
 C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
 C;Accession: S60917; S63237; S65119
 R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.
 submitted to the EMBL Data Library, October 1995
 A;Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV fr
 A;Reference number: S60909
 A;Accession: S60917
 A;Molecule type: DNA
 A;Residues: 1-350 <SEN>
 A;Cross-references: EMBL:X92494; NID:gl045236; PIDN:CAA63233.1; PID:gl045245

R;Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.
 submitted to the Protein Sequence Database, April 1996
 A;Reference number: S63235
 A;Accession: S63237

A;Molecule type: DNA
 A;Residues: 1-350 <SEW>
 A;Cross-references: EMBL:Z71540; NID:gl302320; PIDN:CAA96171.1; PID:gl302321; MIPS:YNI
 A;Experimental source: strain S288C
 R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.
 Yeast 12, 505-514, 1996
 A;Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from
 A;Reference number: S65111; MUID:96310631; PMID:8740425
 A;Accession: S65119
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-350 <SEF>
 A;Cross-references: EMBL:X92494; NID:gl045236; PIDN:CAA63233.1; PID:gl045245
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
 C;Genetics:

A;Gene: SGD:PDR17
 A;Cross-references: SGD:S0005208
 A;Map position: 14L
 A;Note: YNL264
 C;Superfamily: probable membrane protein YNL264c; cellular retinaldehyde-binding prote
 C;Keywords: transmembrane protein
 F;96-291/Domain: cellular retinaldehyde-binding protein homology <CRB>
 F;239-255/Domain: transmembrane #status predicted <TM>

Query Match 87.1%; Score 27; DB 2; Length 350;
 Best Local Similarity 50.0%; Pred. No. 6.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 | | | | |
 Db 249 WFAWAF 254

RESULT 39

S63197
 hypothetical protein YNL231c - yeast (Saccharomyces cerevisiae)
 N;Alternate names: hypothetical protein N1158
 C;Species: Saccharomyces cerevisiae
 C;Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 19-Apr-2002
 C;Accession: S63197; S67367; S72085
 R;Pandolfo, D.; De Antoni, A.; Lanfranchi, G.; Valle, G.
 submitted to the Protein Sequence Database, April 1996
 A;Reference number: S63188
 A;Accession: S63197
 A;Molecule type: DNA
 A;Residues: 1-351 <PAN>
 A;Cross-references: EMBL:Z71507; NID:gl302256; PIDN:CAA96136.1; PID:gl302257; MIPS:YNI
 A;Experimental source: strain S288C
 R;Pandolfo, D.; de Antoni, A.; Lanfranchi, G.; Valle, G.
 submitted to the EMBL Data Library, February 1996
 A;Description: DNA sequence of cosmid 14-5 from chromosome XIV.
 A;Reference number: S67355
 A;Accession: S67367
 A;Molecule type: DNA
 A;Residues: 1-351 <PAW>
 A;Cross-references: EMBL:Z69381; NID:gl183970; PIDN:CAA93367.1; PID:gl183983
 R;Pandolfo, D.; de Antoni, A.; Lanfranchi, G.; Valle, G.
 Yeast 12, 1071-1076, 1996
 A;Title: The DNA sequence of cosmid 14-5 from chromosome XIV reveals 21 open reading f
 A;Reference number: S72073; MUID:97051596; PMID:8896273
 A;Accession: S72085
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-351 <PAF>
 A;Cross-references: EMBL:Z69381; NID:gl183970; PIDN:CAA93367.1; PID:gl183983
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
 C;Genetics:
 A;Gene: SGD:PDR16
 A;Cross-references: SGD:S0005175

A;Map position: 14L

A;Note: YN231c

C;Superfamily: probable membrane protein YN264c; cellular retinaldehyde-binding protein F;91-289/Domain: cellular retinaldehyde-binding protein homology <CRB>

Query Match 87.1%; Score 27; DB 2; Length 351;

Best Local Similarity 50.0%; Pred. No. 6.5e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 247 WLAWTF 252

RESULT 40

A99272

inorganic phosphate transporter SS01183 [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C;Accession: A99272

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: A99272

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-357 <KUR>

A;Cross-references: GB:AE006641; NID:g13814376; PIDN:AAK41432.1; GSPDB:GN00155

C;Genetics:

A;Gene: SS01183

Query Match 87.1%; Score 27; DB 2; Length 357;

Best Local Similarity 50.0%; Pred. No. 6.6e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 328 WASWDF 333

RESULT 41

T33807

hypothetical protein W07B6.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C;Accession: T33807

R;Latreille, P.; Wamsley, P.

submitted to the EMBL Data Library, November 1998

A;Description: The sequence of C. elegans cosmid W07B6.

A;Reference number: Z21414

A;Accession: T33807

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-381 <LAT>

A;Cross-references: EMBL:AF106576; PIDN:AAK78178.1; GSPDB:GN00020; CESP:W07B6.3

A;Experimental source: strain Bristol N2; clone W07B6

C;Genetics:

A;Map position: 2

A;Introns: 27/3; 49/2; 87/3; 149/3; 204/3; 275/3; 336/2

Query Match 87.1%; Score 27; DB 2; Length 381;

Best Local Similarity 50.0%; Pred. No. 7e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 336 WSFWSF 341

RESULT 42

JC7798

vacuole membrane protein 1 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 02-Apr-2002 #sequence_revision 02-Apr-2002 #text_change 02-Apr-2002

C;Accession: JC7798

R;Duseti, N.J.; Jiang, Y.; Vaccaro, M.I.; Tomasini, R.; Samir, A.A.; Calvo, E.L.; Rop-

biochem. Biophys. Res. Commun. 290, 641-649, 2002

A;Title: Cloning and expression of the rat vacuole membrane protein 1 (VMP1), a new ge-

A;Reference number: JC7798

A;Contents: Pancreas

A;Accession: JC7798

A;Molecule type: mRNA

A;Residues: 1-406 <DUS>

C;Comment: This protein is a transmembrane protein located at the Golgi apparatus and

cell death.

C;Genetics:

A;Gene: vmp1

A;Map position: Golgi apparatus and the endoplasmic reticulum area

C;Keywords: transmembrane protein

Query Match 87.1%; Score 27; DB 2; Length 406;

Best Local Similarity 50.0%; Pred. No. 7.3e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 363 WLSWTF 368

RESULT 43

A49424

patterning protein Sonic hedgehog precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000

C;Accession: A49424

R;Riddle, R.D.; Johnson, R.L.; Laufer, E.; Tabin, C.

Cell 75, 1401-1416, 1993

A;Title: Sonic hedgehog mediates the polarizing activity of the ZPA.

A;Reference number: A49424; MUID:94094333; PMID:8289518

A;Accession: A49424

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-425 <RID>

A;Cross-references: GB:L28099; NID:g453526; PIDN:AAA72428.1; PID:g453527

C;Superfamily: sonic hedgehog protein

F;1-26/Domain: signal sequence #status predicted <SIG>

Query Match 87.1%; Score 27; DB 2; Length 425;

Best Local Similarity 50.0%; Pred. No. 7.6e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 361 WAHWF 366

RESULT 44

T20821

hypothetical protein F13B12.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T20821

R;Wild, A.

submitted to the EMBL Data Library, April 1996

A;Reference number: Z19329

A;Accession: T20821

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-431 <WIL>

A;Cross-references: EMBL:Z70683; PIDN:CAA94591.1; GSPDB:GN00022; CESP:F13B12.2

A;Experimental source: clone F13B12

C;Genetics:

A:Gene: CBSP:FL13B12.2

A:Map position: 4

A:Introns: 27/2; 68/1; 135/1; 166/2; 189/3; 233/2; 345/1; 381/2; 400/3

Query Match 87.1%; Score 27; DB 2; Length 431;

Best Local Similarity 50.0%; Pred. No. 7.7e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6

Db 303 WASWGF 308

RESULT 45

hypothetical protein PA3464 [imported] - Pseudomonas aeruginosa (strain PAO1)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: G83212

R:Stover, C.K.; Phan, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: G83212

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-441 <STO>

A:Cross-references: GB:AE004767; GB:AE004091; NID:g9949604; PIDN:AAG06852.1; GSPDB:GN001

A:Experimental source: strain PAO1

C:Genetics:

A:Gene: PA3464

Query Match 87.1%; Score 27; DB 2; Length 441;

Best Local Similarity 50.0%; Pred. No. 7.8e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6

Db 145 WTARWF 150

RESULT 46

AG0842

multidrug resistance protein B [imported] - Salmonella enterica subsp. enterica serovar

C:Species: Salmonella enterica subsp. enterica serovar typhi

A:Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AG0842

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomsen, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

; S.; Mout, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AG0842

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-512 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAP05926.1; PID:g16503897; GSPDB:GN00176

C:Genetics:

A:Gene: STV2941

C:Superfamily: lincomycin-resistance protein lmrB

Query Match 87.1%; Score 27; DB 2; Length 512;

Best Local Similarity 50.0%; Pred. No. 8.8e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6

Db 350 WRATWF 355

RESULT 47

G70556

probable trpE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: G70556

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genom

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: G70556

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-516 <COL>

A:Cross-references: GB:Z95554; GB:AL123456; NID:g3261771; PIDN:CAB08903.1; PID:g32617

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: trpE

C:Superfamily: anthranilate synthase component I

Query Match 87.1%; Score 27; DB 2; Length 516;

Best Local Similarity 50.0%; Pred. No. 8.8e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6

Db 63 WSRWSP 68

RESULT 48

T45254

probable anthranilate synthase [EC 4.1.3.27] component I [imported] - Mycobacterium l

C:Species: Mycobacterium leprae

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C:Accession: T45254

R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, March 1999

A:Reference number: Z22949

A:Accession: T45254

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-529 <JAM>

A:Cross-references: EMBL:AL049913; PIDN:CAB43177.1

A:Experimental source: cosmid B1610

C:Genetics:

A:Note: trpE

C:Superfamily: anthranilate synthase component I

C:Keywords: carbon-carbon lyase; oxo-acid-lyase

Query Match 87.1%; Score 27; DB 2; Length 529;

Best Local Similarity 50.0%; Pred. No. 9e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6

Db 63 WSRWSP 68

RESULT 49

A87596

hypothetical protein CC2801 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: A87596

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.B.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko,

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: A87596

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-628 <STO>

A;Cross-references: GB:AE005673; NID:gl3424403; PIDN:AAK24765.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC2801

Query Match

Best Local Similarity 87.1%; Score 27; DB 2; Length 628;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXP 6

Db 348 WQAWTF 353

RESULT 50

A48646

amine oxidase (copper-containing) (EC 1.4.3.6) - *Arthrobacter* sp. (strain Pl)

C;Species: *Arthrobacter* sp.

C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 29-Sep-1999

C;Accession: A48646

R;Zhang, X.; Fuller, J.H.; McIntire, W.S.

J. Bacteriol. 175, 5617-5627, 1993

A;Title: Cloning, sequencing, expression, and regulation of the structural gene for the

hyolotroph.

A;Reference number: A48646; MUID:93374858; PMID:8365046

A;Accession: A48646

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-648 <ZHA>

A;Cross-references: GB:L12990; NID:g289152; PIDN:AAA22074.1; PID:g289153

C;Superfamily: amine oxidase (copper-containing)

C;Keywords: oxidoreductase

Query Match

Best Local Similarity 87.1%; Score 27; DB 2; Length 648;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXP 6

Db 243 WADWSF 248

Search completed: June 10, 2004, 10:51:23

Job time : 18.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:36:54 ; Search time 8 Seconds
(without alignments)
39.053 Million cell updates/sec

Title: US-09-912-414-9

Perfect score: 31

Sequence: 1 WXXWXP 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : SwisProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	28	90.3	372	1 Y906 MYCTU	Q10562 mycobacteri
2	28	90.3	647	1 ACSA CAUCR	Q9a210 caulobacter
3	28	90.3	766	1 STB6 YEAST	P36085 saccharomyc
4	27	87.1	129	1 YHM2 YEAST	P38857 saccharomyc
5	27	87.1	138	1 YA91 MYCPN	P75602 mycoplasma
6	27	87.1	139	1 YE13 MYCPN	Q9exd6 mycoplasma
7	27	87.1	146	1 YG56 YEAST	P53311 saccharomyc
8	27	87.1	155	1 HOPD ECOLI	O68932 escherichia
9	27	87.1	157	1 YE63 MYCPN	P75320 mycoplasma
10	27	87.1	172	1 YB03 MYCPN	P75566 mycoplasma
11	27	87.1	176	1 VPI Bpp2	P26701 bacterioph
12	27	87.1	229	1 ARTM ARTSA	P17720 artemia sal
13	27	87.1	245	1 YGAZ ECOLI	P76630 escherichia
14	27	87.1	254	1 RCBE RHOC	P19056 rhodobacter
15	27	87.1	350	1 YN04 YEAST	P53844 saccharomyc
16	27	87.1	351	1 YN11 YEAST	P53860 saccharomyc
17	27	87.1	416	1 TWHB BRARE	Q90419 brachydanio
18	27	87.1	425	1 SHH CHICK	Q91035 gallus gall
19	27	87.1	432	1 SHH CYNPY	Q90385 cynops pyrr
20	27	87.1	516	1 TRPE MYCTU	O06127 mycobacteri
21	27	87.1	529	1 TRPE MYCTU	Q9x7c5 mycobacteri
22	27	87.1	648	1 AMO1 ARTS1	Q07123 arthrobacte
23	27	87.1	648	1 AMO2 ARTS1	Q07123 arthrobacte
24	27	87.1	769	1 ITB8 HUMAN	P26012 homo sapien
25	27	87.1	1783	1 RAA3 CHLRE	Q9fec4 chlamydomon
26	26	83.9	79	1 YKFF ECOLI	P75677 escherichia
27	26	83.9	90	1 YPJI ECOLI	P58095 escherichia
28	26	83.9	101	1 YG50 MYCPN	P75147 mycoplasma
29	26	83.9	116	1 HV05 CARAU	P19181 carassius a
30	26	83.9	126	1 CRCE RALSO	O8x2r2 ralstonia s
31	26	83.9	128	1 MERT STRAU	P08656 staphylococ
32	26	83.9	147	1 Y565 METJA	Q57985 methanococc
33	26	83.9	161	1 YVGO BACSU	O32211 bacillus su

34	26	83.9	171	1 Y428 MYCGE	P47667 mycoplasma
35	26	83.9	182	1 YG37 BPMU	Q9t1v8 bacterioph
36	26	83.9	196	1 YOR8 YEAST	P49889 saccharomyc
37	26	83.9	200	1 YH01 MYCTU	P71669 mycobacteri
38	26	83.9	206	1 YN9F ECOLI	P76226 escherichia
39	26	83.9	213	1 IF4E XENLA	P48597 xenopus lae
40	26	83.9	215	1 IF4E APLCA	O77210 aplysia cal
41	26	83.9	217	1 IF4E BOVIN	Q9n0t5 bos taurus
42	26	83.9	217	1 IF4E HUMAN	P07330 homo sapien
43	26	83.9	217	1 IF4E MOUSE	P20415 mus musculu
44	26	83.9	217	1 IF4E RABIT	P29338 oryctolagus
45	26	83.9	226	1 YC07 METJA	Q57604 methanococc
46	26	83.9	234	1 Y523 METJA	Q59643 methanococc
47	26	83.9	234	1 YE99 ARCFU	O28773 archaeglob
48	26	83.9	251	1 Y456 MYCPU	Q98qb0 mycoplasma
49	26	83.9	254	1 OXAA CLOAB	Q97cw0 clostridium
50	26	83.9	256	1 Y090 MYCTU	P10887 mycobacteri
51	26	83.9	257	1 FOL1 HUMAN	P15328 homo sapien
52	26	83.9	301	1 YIJE ECOLI	P32667 escherichia
53	26	83.9	305	1 Y0J4 CABEL	P34627 caenorhabdi
54	26	83.9	314	1 TYSY LISIN	Q22ad4 listeria in
55	26	83.9	314	1 TYSY LISMO	Q9y626 listeria mo
56	26	83.9	316	1 TYSY LACCA	P00469 lactobacill
57	26	83.9	316	1 TYSY LACPL	Q8bw05 lactobacill
58	26	83.9	317	1 TYSY OCEIH	Q8egq0 oceanobacil
59	26	83.9	318	1 TYSY BACAA	Q81e23 bacillus an
60	26	83.9	318	1 TYSY BACCR	Q81e23 bacillus ce
61	26	83.9	318	1 TYSY STAAH	Q99a61 staphylococ
62	26	83.9	318	1 TYSY STAEH	P13954 staphylococ
63	26	83.9	320	1 ASFG SPOPR	O02467 spodoptera
64	26	83.9	321	1 SAPB ECOLI	P47623 escherichia
65	26	83.9	321	1 SAPB SALTY	P36668 salmonella
66	26	83.9	323	1 SBCF ECOLI	P19674 escherichia
67	26	83.9	324	1 YULA BACSU	O34428 bacillus su
68	26	83.9	341	1 YEJE ECOLI	P33915 escherichia
69	26	83.9	348	1 GAL7 ECOLI	P09148 escherichia
70	26	83.9	348	1 GAL7 SALTY	P22714 salmonella
71	26	83.9	350	1 GAL7 DROME	Q9yma2 drosophila
72	26	83.9	358	1 MTC1 CAUCR	Q45971 caulobacter
73	26	83.9	365	1 VNS2 AHSV9	P27279 african hor
74	26	83.9	376	1 MTS1 RHIME	O30569 rhizobium m
75	26	83.9	377	1 MTB1 BRUAB	O30570 bruceella ab
76	26	83.9	379	1 YISP BACCS	P29156 bacillus cl
77	26	83.9	381	1 CYB CHLRE	P23662 chlamydomon
78	26	83.9	381	1 CYE CHLSM	P23663 chlamydomon
79	26	83.9	381	1 G392 DROME	P58957 drosophila
80	26	83.9	381	1 GAL7 CRYNE	P40908 cryptococcu
81	26	83.9	382	1 YCAD ECOLI	P21503 escherichia
82	26	83.9	400	1 ASSY THETH	P59846 thermus the
83	26	83.9	403	1 ASSY LEPIN	Q8eyp7 leptospira
84	26	83.9	404	1 YDIM ECOLI	P76197 escherichia
85	26	83.9	405	1 ASSY RHIME	Q92173 rhizobium m
86	26	83.9	406	1 ASSY BRUSE	Q8yck8 bruceella me
87	26	83.9	406	1 ASSY BRUSU	Q8g376 bruceella su
88	26	83.9	407	1 ASSY AGRT5	Q8uc31 agrobacteri
89	26	83.9	407	1 ASSY RHIL0	Q98e81 rhizobium l
90	26	83.9	409	1 ASSY BUCBP	P59412 buchnera ap
91	26	83.9	412	1 YWFA BACSU	P39637 bacillus su
92	26	83.9	426	1 ENO HELPU	Q9zms6 helicobacte
93	26	83.9	426	1 ENO HELPY	P48285 helicobacte
94	26	83.9	463	1 NU4M STRPU	P15551 strongyloce
95	26	83.9	464	1 NU4M PARLI	P12775 paracentrot
96	26	83.9	477	1 GLGA STRPN	Q97qs5 streptococc
97	26	83.9	482	1 NF31 NAEFO	P42661 naegleria f
98	26	83.9	491	1 Y298 MYCGA	P33276 mycoplasma
99	26	83.9	496	1 REPR STRAG	P18629 streptococc
100	26	83.9	496	1 REPS STRPY	P14752 streptococc
101	26	83.9	508	1 MLO3 ARATH	Q94kb9 arabidopsis
102	26	83.9	521	1 YAV8 SCHPO	Q10177 schizosacch
103	26	83.9	551	1 AGLA RHIME	Q923r8 rhizobium m
104	26	83.9	572	1 YEN1 SCHPO	Q42885 schizosacch
105	26	83.9	576	1 YNSM ANOQU	P33510 anopheles q
106	26	83.9	577	1 YKG6 CABEL	P46556 caenorhabdi

107	26	83.9	578	1	AC22_STRCO	P45105 streptomyce	180	25	80.6	108	1	NIGM_BOVIN	Q02374 bos taurus
108	26	83.9	580	1	P69_MYCHR	P15362 mycoplasma	181	25	80.6	144	1	HV26_MOUSE	P01795 mus musculus
109	26	83.9	609	1	GUP2_YEAST	Q08929 saccharomyc	182	25	80.6	145	1	RL32_AERPE	Q9y52 aeropyrum p
110	26	83.9	634	1	GLGB_LACPL	Q09011 lactobacill	183	25	80.6	164	1	LSPA_ECO57	Q8xa8 escherichia
111	26	83.9	658	1	S282_HUMAN	O43868 homo sapien	184	25	80.6	164	1	LSPA_ECOL6	Q8flb6 escherichia
112	26	83.9	666	1	AMO_LENCU	P49252 lens culina	185	25	80.6	164	1	LSPA_ECOL1	P00804 escherichia
113	26	83.9	666	1	YA35_MYCPN	P75079 mycoplasma	186	25	80.6	165	1	DSBB_VIBAL	Q56578 vibrio algi
114	26	83.9	672	1	YA42_MYCPN	P75072 mycoplasma	187	25	80.6	165	1	LSPA_ENTAE	P13514 enterobacte
115	26	83.9	674	1	AMO_PEA	Q43077 pisum sativ	188	25	80.6	165	1	YRHI_RHOSH	Q32229 rhodobacte
116	26	83.9	742	1	CD44_HUMAN	P16070 h cd44 anti	189	25	80.6	166	1	LSPA_SALTI	Q8z9n1 salmonella
117	26	83.9	750	1	CTPB_MYCLE	P46840 mycobacteri	190	25	80.6	166	1	LSPA_SALTY	Q8zr79 salmoneilla
118	26	83.9	752	1	CTPB_MYCBO	P59947 mycobacteri	191	25	80.6	166	1	Y18K_SSVI	P20209 xylella fas
119	26	83.9	752	1	CTPB_MYCTU	Q10877 mycobacteri	192	25	80.6	167	1	LSPA_XYLFA	Q9paes xylella fas
120	26	83.9	761	1	CTPA_MYCTU	Q10876 mycobacteri	193	25	80.6	167	1	LSPA_XYLFT	Q87b6 xylella fas
121	26	83.9	771	1	TBE8_HUMAN	Q9ncn3 homo sapien	194	25	80.6	168	1	LSPA_VIBVU	Q8des8 vibrio vuln
122	26	83.9	778	1	CD44_MOUSE	P15379 mus musculu	195	25	80.6	169	1	LSPA_PSEAE	Q9hvm5 pseudomonas
123	26	83.9	780	1	CTPA_MYCLE	P46839 mycobacteri	196	25	80.6	169	1	LSPA_VIBPA	Q87s89 vibrio para
124	26	83.9	794	1	VTTB_BPT7	P03747 bacterioph	197	25	80.6	169	1	LSPA_YERPE	Q8zi19 yersinia pe
125	26	83.9	799	1	SYL_MYCPE	Q8w18 mycoplasma	198	25	80.6	170	1	LSPA_PSEFL	P17942 pseudomonas
126	26	83.9	801	1	FTSK_PSESM	Q87z65 pseudomonas	199	25	80.6	171	1	LSPA_PSEPK	Q88q91 pseudomonas
127	26	83.9	801	1	FTSK_PSESY	Q93ul1 pseudomonas	200	25	80.6	171	1	LSPA_VIBSCH	Q9ku46 vibrio chol
128	26	83.9	802	1	SYL_BACAA	O81kk6 bacillus an	201	25	80.6	172	1	DSBB_VIBVU	P59347 vibrio vuln
129	26	83.9	802	1	SYL_BACCR	O81kk6 bacillus ce	202	25	80.6	173	1	LSPA_PSESM	Q88963 pseudomonas
130	26	83.9	803	1	SYL_LTSIN	Q92az9 listeria in	203	25	80.6	173	1	LSPA_PSESM	Q9ung2 homo sapien
131	26	83.9	803	1	SYL_LTSIM	Q92az9 listeria in	204	25	80.6	177	1	TN18_HUMAN	Q9hd36 homo sapien
132	26	83.9	804	1	SYL_BACSU	Q8v6m4 listeria mo	205	25	80.6	194	1	BCLB_HUMAN	P37616 escherichia
133	26	83.9	804	1	SYL_OCEIH	P36430 bacillus su	206	25	80.6	208	1	YHEN_ECOLI	Q37616 escherichia
134	26	83.9	804	1	SYL_STAAM	Q8ep12 oceanobacil	207	25	80.6	210	1	SSH5_YEAST	Q03446 saccharomyc
135	26	83.9	806	1	SYL_BACHD	Q997a8 staphylococ	208	25	80.6	219	1	ENGB_METJA	Q57768 methanococc
136	26	83.9	806	1	SYL_UREPA	Q9k7s8 bacillus ha	209	25	80.6	237	1	YNP5_CAEEL	P34558 caenorhabdi
137	26	83.9	807	1	SYL_STAAP	Q9pqc0 ureaplasma	210	25	80.6	239	1	LPXH_PASMU	Q3cpe3 pasteurella
138	26	83.9	808	1	SYL_LACPL	O8w17 staphylococ	211	25	80.6	242	1	LPXH_VIBPA	Q87gk1 vibrio para
139	26	83.9	811	1	FTSK_PSEAE	O8x9a9 lactobacill	212	25	80.6	243	1	YKJA_BACSU	P49853 bacillus su
140	26	83.9	812	1	SYL_CLOAB	Q910m3 pseudomonas	213	25	80.6	245	1	TATC_CAMTE	Q9phE8 campylobact
141	26	83.9	812	1	SYL_CLOTE	Q89lv6 clostridium	214	25	80.6	249	1	TATC_HELPJ	Q9zm59 helicobacte
142	26	83.9	814	1	SYL_STAEP	Q89lv6 clostridium	215	25	80.6	254	1	YABI_ECOLI	Q25701 helicobacte
143	26	83.9	815	1	SYL_NACFA	Q8cmu8 staphylococ	216	25	80.6	257	1	Y405_UREPA	Q30149 escherichia
144	26	83.9	816	1	SYL_CLOPE	Q9gl97 macaca fasc	217	25	80.6	260	1	COBS_SYNV3	Q57174 synechocyst
145	26	83.9	821	1	CAN3_HUMAN	O8xm18 clostridium	218	25	80.6	271	1	SC5D_TOBAC	Q9zt29 nicotiana t
146	26	83.9	821	1	CAN3_MOUSE	P20807 homo sapien	219	25	80.6	276	1	GLPG_ECOLI	P09391 escherichia
147	26	83.9	821	1	CAN3_FIG	Q64691 mus musculu	220	25	80.6	279	1	SC51_ARATH	Q9m883 arabidopsis
148	26	83.9	821	1	CAN3_RAT	P43368 sus scrofa	221	25	80.6	281	1	SC51_ARATH	Q92088 arabidopsis
149	26	83.9	824	1	SYL_THEMA	Q9wy15 thermotoga	222	25	80.6	306	1	YK26_YEAST	P36139 saccharomyc
150	26	83.9	829	1	SYL_LACLA	Q9hb6 lactococcus	223	25	80.6	306	1	CDSA_MYCTU	Q10807 m phosphati
151	26	83.9	833	1	SYL_STR3	Q8e2v2 streptococc	224	25	80.6	307	1	TR41_HUMAN	P59536 homo sapien
152	26	83.9	833	1	SYL_STRMU	Q8e2v2 streptococc	225	25	80.6	307	1	TR59_HUMAN	P59550 homo sapien
153	26	83.9	833	1	SYL_STRP3	Q8d885 streptococc	226	25	80.6	308	1	T2RC_MOUSE	P59532 mus musculu
154	26	83.9	833	1	SYL_STRP8	Q8d881 streptococc	227	25	80.6	308	1	T2RC_RAT	Q9jke7 rattus norv
155	26	83.9	833	1	SYL_STRPN	Q8p2t2 streptococc	228	25	80.6	318	1	Y501_SYNV3	Q55487 synechocyst
156	26	83.9	833	1	SYL_STRPY	Q97as0 streptococc	229	25	80.6	345	1	YA03_TREPA	O83968 treponema p
157	26	83.9	833	1	SYL_STRR6	Q8drb6 streptococc	230	25	80.6	349	1	GAL7_HAGIN	P31764 haemophilus
158	26	83.9	834	1	FTSK_PSEPK	O8sf88 pseudomonas	231	25	80.6	349	1	HMEN_ARTSF	Q05640 artemia san
159	26	83.9	840	1	SYL_FORBU	O51267 borrelia bu	232	25	80.6	349	1	LL8A_RAT	P70612 rattus norv
160	26	83.9	857	1	SYL_SYNEL	Q8dh61 synechococc	233	25	80.6	354	1	VNS2_ETV17	P33473 bluetongue
161	26	83.9	869	1	SYL_SYNV3	P73274 synechocyst	234	25	80.6	357	1	VNS2_ETV10	P33065 bluetongue
162	26	83.9	878	1	SYL_TREPA	O83595 treponema p	235	25	80.6	357	1	VNS2_BTW1S	P32932 bluetongue
163	26	83.9	952	1	SYL_CORGL	Q8alc4 corynebacte	236	25	80.6	357	1	VNS2_BTW1X	P10350 bluetongue
164	26	83.9	957	1	SYL_COREF	Q8flm0 corynebacte	237	25	80.6	358	1	CKR3_CAVPO	Q92213 cavia porce
165	26	83.9	966	1	SYL_STRCO	Q9rd15 streptomyce	238	25	80.6	359	1	CKR3_MOUSE	P51678 mus musculu
166	26	83.9	969	1	SYL_MYCTU	P71698 mycobacteri	239	25	80.6	359	1	CKR3_RAT	Q54814 rattus norv
167	26	83.9	972	1	SYL_MYCLE	Q50192 mycobacteri	240	25	80.6	365	1	GAL7_YEAST	P08431 saccharomyc
168	26	83.9	999	1	MGR1_CAEEL	Q9u630 caenorhabdi	241	25	80.6	369	1	GAL7_SCHPO	Q9hdus schizosacch
169	26	83.9	1008	1	YJX4_SCHPO	Q9uun9 schizosacch	242	25	80.6	369	1	Y316_MYGGE	P47558 mycoplasma
170	26	83.9	1021	1	CPOM_NEUCR	P33537 neurospora	243	25	80.6	370	1	GAL7_KJULA	P09580 kluyveromyc
171	26	83.9	1183	1	CNA_STAAN	Q53654 staphylococ	244	25	80.6	370	1	TAM2_HUMAN	Q15035 homo sapien
172	26	83.9	1386	1	Y064_MYCPN	P75613 mycoplasma	245	25	80.6	376	1	CYB_FIABE	Q9253 plasmodium
173	26	83.9	3744	1	TRAI_YEAST	P38811 saccharomyc	246	25	80.6	376	1	CYB_PLACH	Q9253 plasmodium
174	25	80.6	65	1	LHA2_ECTHL	P80103 ectothiorho	247	25	80.6	376	1	CYB_PLACH	Q02768 plasmodium
175	25	80.6	78	1	VG9_SPV1R	P15900 spiroplasma	248	25	80.6	376	1	TRUC_ERWCA	Q47417 erwina car
176	25	80.6	84	1	ACHA_CRORS	P54248 crocidura r	249	25	80.6	376	1	TRUC_ERWCA	P27280 epizootic h
177	25	80.6	84	1	ACHA_FELCA	P54250 felis silve	250	25	80.6	379	1	GAL7_MOUSE	Q03249 mus musculu
178	25	80.6	105	1	NIGM_HUMAN	Q95178 homo sapien	251	25	80.6	379	1	GAL7_RAT	P43424 rattus norv
179	25	80.6	106	1	YQCC_HABIN	Q57152 haemophilus	252	25	80.6	380	1	CYB_RANCA	P16674 rana catesb

253	25	80.6	382	1	CYB PLAYS	063696	plasmodium	325	25	80.6	873	1	SVL_PSEAE	Q9hx33	pseudomonas
254	25	80.6	382	1	GAL7 FRIRE	Q96uil	trichoderma	327	25	80.6	876	1	SVL_NEIMB	Q9jxt2	neisseria m
255	25	80.6	388	1	NH16_CABEL	Q27521	caenorhabdi	329	25	80.6	906	1	RIRI_HCMVA	Q91w39	human cytom
256	25	80.6	402	1	ASSY DEIRA	Q9rwj4	deinococcus	329	25	80.6	910	1	SVL_NEIMA	Q91w39	neisseria m
257	25	80.6	405	1	DBR1 YEAST	P24309	saccharomyc	330	25	80.6	918	1	PEP3 YEAST	P27801	lycopersico
258	25	80.6	407	1	NH86_CABEL	Q965w2	caenorhabdi	331	25	80.6	960	1	VP41 LYCES	P93231	saccharomyc
259	25	80.6	408	1	ASSY CAUCR	Q9abul	caulobacter	332	25	80.6	969	1	DPOM NEUTN	P33538	neurospora
260	25	80.6	420	1	GLGC RHIME	Q92ml3	rhizobium m	333	25	80.6	972	1	CPAL_BACCI	P34286	bacillus ci
261	25	80.6	424	1	FD6C SOYEN	P48628	glycine max	334	25	80.6	976	1	VP41_ARATH	P93043	arabidopsis
262	25	80.6	433	1	YBBY ECOLI	P77328	escherichia	335	25	80.6	988	1	TERT_SCHPO	O13339	schizosacch
263	25	80.6	436	1	ACHX ONCVO	P54247	onchocerca	336	25	80.6	1010	1	ALA3_OREMO	P58312	oreochromis
264	25	80.6	439	1	Y412 ARATH	O46558	arabidopsis	337	25	80.6	1037	1	ATC3 SCHPO	P22189	schizosacch
265	25	80.6	443	1	FD6C BRANA	P48627	brassica na	338	25	80.6	1062	1	SUM1 YEAST	P46676	saccharomyc
266	25	80.6	448	1	FD6C ARATH	P46312	arabidopsis	339	25	80.6	1063	1	DPOM_CLAPU	P22373	claviceps p
267	25	80.6	456	1	YCH3_ASTLO	P14761	astasia lon	340	25	80.6	1197	1	DPOM_PODAN	O01529	podospira a
268	25	80.6	457	1	ACHA BOVIN	P02709	bos taurus	341	25	80.6	2329	1	YLO6_CABEL	P34369	caenorhabdi
269	25	80.6	457	1	ACHA MOUSE	P04756	mus musculus	342	25	80.6	2936	1	NBEA MOUSE	Q9epn1	mus musculus
270	25	80.6	457	1	ACHA RAT	P25108	rattus norv	343	25	80.6	2946	1	NBEA HUMAN	Q8nfp9	homo sapien
271	25	80.6	469	1	GABP BACSU	P46349	bacillus su	344	24	77.4	129	1	KV3H HUMAN	P04207	homo sapien
272	25	80.6	475	1	ABPE STRCO	O54161	streptomyce	345	24	77.4	140	1	LYSA_DROME	P37157	drosophila
273	25	80.6	475	1	ABPE STRLI	P96463	streptomyce	346	24	77.4	140	1	LYSB_DROME	Q08594	drosophila
274	25	80.6	478	1	MLJ3 ARATH	Q54MB2	arabidopsis	347	24	77.4	141	1	LYSE_DROME	P29615	drosophila
275	25	80.6	479	1	CATA PSEPU	Q59714	pseudomonas	348	24	77.4	142	1	LYSX_DROME	P37161	drosophila
276	25	80.6	483	1	COA2 BPFP3	P03624	bacterioph	349	24	77.4	142	1	DSBB_VIBPA	Q87n03	vibrio para
277	25	80.6	496	1	ML15 ARATH	O80580	arabidopsis	350	24	77.4	178	1	RP28 BPSP1	P03048	bacterioph
278	25	80.6	499	1	MEP2 YEAST	P41948	saccharomyc	351	24	77.4	220	1	FLIR_BUCAL	P57186	buchnera ap
279	25	80.6	509	1	GLPK PSESM	O87x10	pseudomonas	352	24	77.4	258	1	IBAL_HUMAN	Q14556	homo sapien
280	25	80.6	503	1	GLPK PSETO	O87524	pseudomonas	353	24	77.4	261	1	DET2_ARATH	Q93944	arabidopsis
281	25	80.6	504	1	SIK1 YEAST	Q12460	saccharomyc	354	24	77.4	262	1	EUTT_SALTY	Q9zfv4	salmonella
282	25	80.6	508	1	C983 ARATH	O22203	arabidopsis	355	24	77.4	267	1	GUB RHOMR	P45798	rhodothermu
283	25	80.6	509	1	C982 SOYEN	O48922	glycine max	356	24	77.4	286	1	SURI MOUSE	P09925	mus musculus
284	25	80.6	509	1	YRV4_CABEL	Q27516	caenorhabdi	357	24	77.4	306	1	OPPB_BACSU	P24138	bacillus su
285	25	80.6	511	1	NOPE YEAST	O12499	saccharomyc	358	24	77.4	311	1	Y350_MYCSE	P47592	mycoplasma
286	25	80.6	526	1	ML01 ARATH	O49621	arabidopsis	359	24	77.4	311	1	RDHH HUMAN	Q8tc12	homo sapien
287	25	80.6	529	1	NOPS HUMAN	Q9y2x3	homo sapien	360	24	77.4	318	1	FUT1_RABIT	Q91979	oryctolagus
288	25	80.6	534	1	NOPS RAT	Q9qz86	rattus norv	361	24	77.4	373	1	CYRG BOVIN	Q95118	bos taurus
289	25	80.6	551	1	QCRB MYCLE	P15878	mycobacteri	362	24	77.4	379	1	O59A_DROME	P81923	drosophila
290	25	80.6	560	1	NMB HUMAN	Q14956	homo sapien	363	24	77.4	397	1	ENO_2ALTA	Q9hdt3	alternaria
291	25	80.6	573	1	ALPI YEAST	P38971	saccharomyc	364	24	77.4	438	1	ENO_ASPOR	Q12560	aspergillus
292	25	80.6	587	1	ZDS_TARER	Q9fv46	tagetes ere	365	24	77.4	438	1	ENO_ASPOR	P24040	cladosporiu
293	25	80.6	595	1	S131 HUMAN	Q9bwz2	homo sapien	366	24	77.4	440	1	NUAM CHLRE	P20113	chlamydomon
294	25	80.6	595	1	S131 RAT	Q07782	rattus norv	367	24	77.4	443	1	NH12_CABEL	Q21701	caenorhabdi
295	25	80.6	606	1	Z214 HUMAN	Q9ul59	homo sapien	368	24	77.4	453	1	VIPR MELGA	Q91085	melagris g
296	25	80.6	625	1	TPOR MOUSE	O08351	mus musculus	369	24	77.4	457	1	GABP ECOLI	P25527	escherichia
297	25	80.6	627	1	S134 HUMAN	Q9ukq4	homo sapien	370	24	77.4	466	1	GABP ECOLI	Q94kb1	arabidopsis
298	25	80.6	634	1	SYLA AQUAE	O66680	aquifex aeo	371	24	77.4	485	1	ENT_ENTCO	P81007	enterolobiu
299	25	80.6	641	1	MIA2 HUMAN	O60476	homo sapien	372	24	77.4	509	1	CCBS_MARPO	P36180	marchantia
300	25	80.6	641	1	MIA2 MOUSE	P39098	mus musculus	373	24	77.4	512	1	FEN2 YEAST	P25821	saccharomyc
301	25	80.6	649	1	QOX1 BACSU	P34956	bacillus su	374	24	77.4	547	1	GSPA AERHY	P45754	aeromonas h
302	25	80.6	657	1	CSPI CORGL	O01377	corynebacte	375	24	77.4	552	1	NRPE ECOLI	P32710	escherichia
303	25	80.6	671	1	AM01 ASPNG	O12556	aspergillus	376	24	77.4	554	1	ML14 ARATH	Q94kb1	arabidopsis
304	25	80.6	701	1	YG5L YEAST	P53326	saccharomyc	377	24	77.4	573	1	ML11 ARATH	Q9fi00	arabidopsis
305	25	80.6	702	1	FOX4_SALTY	O56145	salmonella	378	24	77.4	577	1	CCBS_OENBE	Q04648	oenothera b
306	25	80.6	740	1	YCF2_CUSRE	P32033	cuscuta ref	379	24	77.4	581	1	FUR4_SCHPO	Q10279	schizosacch
307	25	80.6	755	1	AMO KLEAB	P49250	klebsiella	380	24	77.4	630	1	YCF2_OENVI	P31569	oenothera v
308	25	80.6	757	1	AMO ECOLI	P46883	escherichia	381	24	77.4	635	1	NRPE_HAEIN	P44944	haemophilus
309	25	80.6	769	1	TMB8 MOUSE	Q9esn3	mus musculus	382	24	77.4	640	1	Y4CD RHISN	P55386	rhizobium s
310	25	80.6	778	1	PMTY SCHPO	O42933	schizosacch	383	24	77.4	647	1	CCMF ECOLI	P33927	escherichia
311	25	80.6	788	1	BCSP_XANAC	P58933	xanthomonas	384	24	77.4	648	1	CCMF_HAEIN	P45037	haemophilus
312	25	80.6	823	1	SYL DEIRA	Q9rf60	deinococcus	385	24	77.4	660	1	CCMF_BRAJA	P45037	bradyrhizob
313	25	80.6	850	1	STB2 YEAST	P46679	saccharomyc	386	24	77.4	660	1	CCMF_PSEFL	P52225	pseudomonas
314	25	80.6	852	1	WS14 HUMAN	Q9np71	homo sapien	387	24	77.4	676	1	CCMF_RHIME	P45404	rhizobium m
315	25	80.6	859	1	SYL_BUCAL	P57519	buchnera ap	388	24	77.4	717	1	YCCS ECOLI	P75870	escherichia
316	25	80.6	859	1	SYL_FUSNN	Q8ric3	fusobacteri	389	24	77.4	721	1	PLSB_CAEEL	Q22949	caenorhabdi
317	25	80.6	859	1	SYL_SHEON	Q8ehp4	shewanella	390	24	77.4	721	1	YCF2_OENPI	P31568	oenothera p
318	25	80.6	860	1	SYL_PASMO	P57923	pasteurella	391	24	77.4	788	1	RBC2_HAEIN	P44408	haemophilus
319	25	80.6	860	1	SYL YERPE	Q8zdf8	yersinia pe	392	24	77.4	799	1	AFSK STRCO	P54741	streptomyce
320	25	80.6	861	1	SYL_BUCAP	Q8k9b9	buchnera ap	393	24	77.4	807	1	AFSK STRGR	P54742	streptomyce
321	25	80.6	861	1	SYL_BUCBP	P59433	buchnera ap	394	24	77.4	982	1	POL_HTLV2	P03363	human t-cel
322	25	80.6	861	1	SYL_HAEIN	P43827	haemophilus	395	24	77.4	1257	1	ERB2 RAT	P06494	rattus norv
323	25	80.6	864	1	WS14 MOUSE	Q99mz3	mus musculus	396	24	77.4	1535	1	LML1_CAEEL	Q18823	caenorhabdi
324	25	80.6	868	1	SYL_PSEPK	Q88dnl	pseudomonas	397	24	77.4	2131	1	YCF2_SPIOL	P08973	spinacia ol
325	25	80.6	868	1	SYL_PSESM	Q87vx3	pseudomonas	398	24	77.4	2216	1	YCF2_EPIVI	P30072	epifagus vl

399	24	77.4	2280	1	YCF2_OENHO	Q9mf2f oenothera h	472	23	74.2	961	1	GPR1_YEAST	Q12361 saccharomyc
400	24	77.4	2280	1	YCF2_TOBAC	P09976 nicotiana t	473	23	74.2	964	1	CTA2_BACCI	P70873 bacillus ci
401	24	77.4	117	1	HV30_HUMAN	P01776 homo sapien	474	23	74.2	974	1	CC15_YEAST	P27636 saccharomyc
402	23	74.2	119	1	HV3M_HUMAN	P01774 homo sapien	475	23	74.2	1041	1	CHS1_CRYNE	Q13356 cryptococcu
403	23	74.2	119	1	HV3N_HUMAN	P01775 homo sapien	476	23	74.2	1110	1	VGLM_INSV	Q11260 impatiens n
404	23	74.2	137	1	VGLL_EBV	P03212 Epstein-Bar	477	23	74.2	1739	1	CHD2_HUMAN	Q14647 homo sapien
405	23	74.2	140	1	LYSS_DROME	P37160 drosophila	478	23	74.2	1807	1	VTA2_XENLA	P18709 xenopus lae
406	23	74.2	162	1	YBJO_ECOLI	P75816 escherichia	479	23	74.2	2298	1	KROS_HUMAN	Q9b1k6 lotus japon
407	23	74.2	184	1	LOLA_HELPF	Q92158 helicobacte	480	23	74.2	2347	1	YCF2_LOTJA	P08922 homo sapien
408	23	74.2	184	1	LOLA_HELPF	Q25474 helicobacte	481	22	71.0	66	1	RPB1_CABER	P35074 caenorhabdi
409	23	74.2	219	1	VIF_SIVAI	Q02841 simian immu	482	22	71.0	86	1	ACBP_CABER	Q01805 caenorhabdi
410	23	74.2	229	1	REGO_BP82	P13870 bacterioph	483	22	71.0	88	1	Y419_NEIMA	Q9jgk5 neisseria m
411	23	74.2	232	1	CLB3_SHEEP	P80943 ovis aries	484	22	71.0	90	1	F883_XYLFT	Q87d06 xylella fas
412	23	74.2	238	1	Y206_CHLMO	Q9plal chlamydia m	485	22	71.0	90	1	YJ08_XYLFA	Q9pc73 xylella fas
413	23	74.2	238	1	H819_CHLTR	Q84826 chlamydia t	486	22	71.0	90	1	HMEN_SCHAM	P14150 schistocerc
414	23	74.2	260	1	HIS6_CAUCR	Q9a229 caulobacte	487	22	71.0	93	1	YQJK_ECOLI	Q47710 escherichia
415	23	74.2	262	1	CAHD_HUMAN	Q8n1q1 homo sapien	488	22	71.0	102	1	THIO_MYCGE	P47370 mycoplasma
416	23	74.2	262	1	CAHD_MOUSE	Q8d6n1 mus musculu	489	22	71.0	103	1	THIO_LISMO	Q9s386 listeria mo
417	23	74.2	294	1	LYCM_STRGL	P25310 streptomyce	490	22	71.0	104	1	THIO_STAAM	Q9s384 staphylococ
418	23	74.2	301	1	YF34_SYNY3	P74220 synechocyst	491	22	71.0	104	1	THIO_STAEP	Q8cp15 staphylococ
419	23	74.2	303	1	REG1_SCHPO	P40389 schizosacch	492	22	71.0	105	1	THIO_ALIAC	P80579 alicyclobac
420	23	74.2	305	1	CAH5_HUMAN	P35218 homo sapien	493	22	71.0	106	1	THIO_CPCM	Q9uw02 coprinus co
421	23	74.2	312	1	YCBK_BACSU	P42243 bacillus su	494	22	71.0	108	1	THI2_CHLTE	Q8ke49 chlorobium
422	23	74.2	314	1	MTB1_HABPA	P29538 haemophilus	495	22	71.0	108	1	THIO_CHLJT	P10472 chlorobium
423	23	74.2	333	1	CLB1_SHEEP	Q28565 ovis aries	496	22	71.0	109	1	THI2_SYNY3	P73263 synechocyst
424	23	74.2	333	1	CLB2_SHEEP	Q29422 ovis aries	497	22	71.0	110	1	THI2_ANASP	P20857 anaebactera
425	23	74.2	341	1	Y534_TREPA	Q83545 treponema p	498	22	71.0	110	1	THIO_STRCO	P52230 streptomyce
426	23	74.2	349	1	ACOB_TRINI	O44390 trichoplusi	499	22	71.0	110	1	THIO_MYCSM	O30974 mycobacteri
427	23	74.2	369	1	Y264_SYNY3	P73879 synechocyst	500	22	71.0	115	1	THIO_MYCTU	P52229 mycobacteri
428	23	74.2	370	1	SERC_METAC	O8n1i1 methanosarc							
429	23	74.2	370	1	SERC_METRA	P52878 methanosarc							
430	23	74.2	370	1	SERC_METWA	Q9pt12 methanosarc							
431	23	74.2	372	1	YC60_MYCTU	Q11058 mycobacteri							
432	23	74.2	377	1	SEN2_YEAST	P16658 saccharomyc							
433	23	74.2	379	1	GAL7_HUMAN	P07902 homo sapien							
434	23	74.2	381	1	FLHB_BUCAP	Q8k9s1 buchera ap							
435	23	74.2	391	1	HERP_HUMAN	Q15011 homo sapien							
436	23	74.2	391	1	HERP_MOUSE	Q9jjk5 mus musculu							
437	23	74.2	396	1	BCR_ECOLI	P28246 escherichia							
438	23	74.2	413	1	INVB_ZYMO	Q60115 zymomonas m							
439	23	74.2	418	1	GATD_ARCFU	Q92380 arcaeaoglob							
440	23	74.2	421	1	EXG_YARLI	Q12725 yarrowia li							
441	23	74.2	431	1	ENO_SHEON	Q9ebro shewanella							
442	23	74.2	440	1	Y485_MYCPN	P75608 mycoplasma							
443	23	74.2	466	1	GUN5_THEFU	Q01786 thermomonos							
444	23	74.2	478	1	SHT3_HUMAN	P46098 homo sapien							
445	23	74.2	483	1	SHT3_RAT	P35563 rattus norv							
446	23	74.2	487	1	SHT3_MOUSE	P23979 mus musculu							
447	23	74.2	490	1	SHT3_CAVPO	O70212 cavia porce							
448	23	74.2	497	1	DTPT_LACHE	O07380 lactobacill							
449	23	74.2	504	1	CAIT_PROSL	P59334 proteus sp.							
450	23	74.2	549	1	QCRB_MYCTU	Q10388 mycobacteri							
451	23	74.2	554	1	HYES_HUMAN	P34913 homo sapien							
452	23	74.2	554	1	HYES_MOUSE	P34914 mus musculu							
453	23	74.2	554	1	HYES_RAT	P80299 rattus norv							
454	23	74.2	609	1	LKHA_RAT	P30349 rattus norv							
455	23	74.2	610	1	LKHA_CAVPO	P19602 cavia porce							
456	23	74.2	610	1	LKHA_HUMAN	P09960 homo sapien							
457	23	74.2	610	1	LKHA_MOUSE	P24527 mus musculu							
458	23	74.2	614	1	YG3F_YEAST	P53283 saccharomyc							
459	23	74.2	658	1	GUN3_FIBSU	P14250 fibrobacter							
460	23	74.2	671	1	LKHA_YEAST	Q10740 saccharomyc							
461	23	74.2	707	1	CAD1_LISMO	P58414 listeria mo							
462	23	74.2	711	1	CAD2_LISMO	Q60048 listeria mo							
463	23	74.2	723	1	CADA_BACPF	P30336 bacillus ps							
464	23	74.2	727	1	CADA_STAUA	P20021 staphylococ							
465	23	74.2	731	1	BGAL_DIAAC	Q00662 dianthus ca							
466	23	74.2	804	1	CADD_STAUA	P37386 staphylococ							
467	23	74.2	822	1	CAN3_BOVIN	P51186 bos taurus							
468	23	74.2	822	1	CAN3_SHEEP	Q9tth8 ovis aries							
469	23	74.2	824	1	SLIP_SALTY	Q9tth8 ovis aries							
470	23	74.2	824	1	SLIP_SALTY	Q9tth8 ovis aries							
471	23	74.2	827	1	GP21_YEAST	P40433 saccharomyc							
472	23	74.2	857	1	VGLB_EBV	P03188 epstein-bar							

ALIGNMENTS

RESULT 1
Y906 MYCTU STANDARD; PRT; 372 AA.
AC Q10562; (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein RV0906/MT0929/MD0930 precursor.
GN RV0906 OR MT0929 OR MTCY31.34 OR MD0930.
OS Mycobacterium tuberculosis, and
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
Hornsbey T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
Delecher A., Uitterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and

laboratory strains.";
J. Bacteriol. 184:5479-5490 (2002).
[3]
SEQUENCE FROM N.A.
SPECIES=M.BOVIS; STRAIN=AF2122/97;
MEDLINE=22709107; PubMed=12788972;
Garnier T., Eighmeier K., Camus J.-C., Medina N., Mansoor H.,
Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
"The complete genome sequence of Mycobacterium bovis.";
Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
-!- SIMILARITY: TO K.PNEUMONIAE ROMA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z73101; CAA97381.1; -;
DR EMBL; AE006979; AAK45176.1; -;
DR EMBL; BX248337; CAD93791.1; -;
DR PIR; G70783; G70783.
DR TIGR; MT0929; -;
DR TubercuList; RV0906; -;
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 372 HYPOTHETICAL PROTEIN
FT FT RV0906/MT0929/MB0930.
FT SQ SEQUENCE 372 AA; 40641 MW; 0A85549D2429D335 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 372;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWXF 6
DB 243 WASWAF 248

RESULT 2
ACSA CAUCR STANDARD; PRT; 647 AA.
ID ACSA CAUCR
AC Q9A210;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Acetyl-coenzyme A synthetase (EC 6.2.1.1) (Acetate--CoA ligase) (Acyl-
DE activating enzyme).
DE ACSA OR CC3581.
GN Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 19089 / CB15;
RC MEDLINE=21173698; PubMed=11259647;
RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
Potočka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
Uttarback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
Salzberg S.L., Venkter J.C., Shapiro L., Fraser C.M.;
"Complete genome sequence of Caulobacter crescentus.";
Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
CC -!- CATALYTIC ACTIVITY: ATP + acetate + COA = AMP + diphosphate +
CC acetyl-CoA.
CC -!- PWM: Acetylated. Deacetylation by the SIR2-homolog deacetylase

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; X75780; CAA53402.1; -;
DR EMBL; Z28072; CAA81909.1; -;
DR PIR; S37894; S37894;
DR Germline; 139828; -;
DR SGD; S0001555; YK072W;
SQ SEQUENCE 766 AA; 88835 MW; 9BCBEA2EE03A9AF1 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 766;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | | | |
Db 745 WSTWTF 750

RESULT 4

YH2 YEAST STANDARD; PRT; 129 AA.

AC P38857; 1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 14.6 kDa protein in REC104-SOL3 intergenic region.
GN YHR162W.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

NCBI_TaxID=4932;

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;

RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
Kucaba T., Hillier L.W., Jier M., Johnston L., Langston J.,
Lacaille P., Louie E.J., Macri C., Mardis E., Meneses S., Mouser L.,
Nhan M., Rifkin L., Riles L., St Peter H., Trevaakis E., Vaughan K.,
Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII.";

RL Science 265:2077-2082(1994).
CC -!- SIMILARITY: BELONGS TO THE UPF0041 (0-44) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; U00027; AAB68009.1; -;
DR PIR; S48902; S48902;
DR Germline; 139480; -;
DR SGD; S0001205; YHR162W

DR InterPro; IPR005336; UPF0041.

DR Pfam; PF03650; UPF0041; 1.

KW Hypothetical protein.

SQ SEQUENCE 129 AA; 14555 MW; 5C1FB3A463DE2A7 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 129;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | | | |
Db 64 WTRWSF 69

RESULT 5

YA91 MYCPN

ID YA91 MYCPN STANDARD; PRT; 138 AA.

AC P75602;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein MPN091 (R02_orf138).

GN MPN091 OR MP064.

OS Mycoplasma pneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

NCBI_TaxID=2104;

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 29342 / M129;

RX MEDLINE=97105885; PubMed=8948633;

RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
Herrmann R.;

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae";

RL Nucleic Acids Res. 24:4420-4449(1996).

CC -!- SIMILARITY: ALMOST IDENTICAL TO M.PNEUMONIAE MPN413 AND MPN463.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AE000008; AAB95711.1; -;

DR PIR; S73390; S73390.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 138 AA; 15763 MW; F52380D428F42709 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 138;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | | | |
Db 36 WSGWSF 41

RESULT 6

YE13 MYCPN

ID YE13 MYCPN STANDARD; PRT; 139 AA.

AC Q9EXD6;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein MPN413 (A05_orf139).

GN MPN413 OR MP426.1.

OS Mycoplasma pneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

NCBI_TaxID=2104;

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 29342 / M129;

RX MEDLINE=97105885; PubMed=8948633;

RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
Herrmann R.;

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae";

RL Nucleic Acids Res. 24:4420-4449(1996).

CC IDENTIFICATION.

```

RX MEDLINE=20411492; PubMed=10954595;
RA Dandekar T., Huynen M., Regula J.T., Ueberle B., Zimmermann C.U.,
RA Andrade M.A., Doerks T., Sanchez-Pulido L., Snel B., Suyama M.,
RA Yuan Y.P., Herrmann R., Bork P.;
RT "Re-annotating the Mycoplasma pneumoniae genome sequence: adding
RT value, function and reading frames.";
RL Nucleic Acids Res. 28:3278-3288(2000).
CC -!- SIMILARITY: ALMOST IDENTICAL TO M.PNEUMONIAE MPN091 AND MPN463.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000041; AAC34747.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 139 AA; 15795 MW; 0CD89C29543B5182 CRC64;

Query Match      87.1%; Score 27; DB 1; Length 139;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
DB 37 WSGWSF 42

RESULT 7
Y056_YEAST
ID YG56 YEAST STANDARD; PRT; 146 AA.
AC P53311;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 16.2 kDa protein in PFK1-TDS4 intergenic region.
GN YGR243W OR G8620.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN SEQUENCE FROM N.A.
RP STRAIN=S288C;
RC MEDLINE=97245298; PubMed=9090057;
RX Guerreiro P., Azevedo D., Barriros T., Rodrigues-Pousada C.;
RT "Sequencing of a 9.9 kb segment on the right arm of yeast chromosome
RT VII reveals four open reading frames, including PFK1, the gene coding
RT for succinyl-CoA synthetase (beta-chain) and two ORFs sharing
RT homology with ORFs of the yeast chromosome VIII.";
RL Yeast 13:275-280(1997).
CC -!- SIMILARITY: BELONGS TO THE UPF0041 (0-44) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z73028; CAA97272.1; -.
DR PIR; S64569; S64569.
DR GerMOnline; 141555; -.
DR SGD; S0003475; YGR243W.
DR InterPro; IPR005336; UPF0041.
DR Pfam; PF03650; UPF0041; 1.
KW Hypothetical protein.
SQ SEQUENCE 146 AA; 16230 MW; E0B13933AB142B4E CRC64;

Query Match      87.1%; Score 27; DB 1; Length 146;

```

```

Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
DB 64 WTRWSF 69

RESULT 8
HOPD_ECOLI
ID HOPD_ECOLI STANDARD; PRT; 155 AA.
AC O68932;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leader peptidase hopd.
GN HOPD.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]_TaxID=562;
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR30;
RA Noorani S.M., Lindahl L., Zengel J.M.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family A24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF058450; AAC14289.1; -.
DR MEROPS; A24.003; -.
DR InterPro; IPR000045; Peptidase A24A.
DR Pfam; PF01478; Peptidase A24; 1.
DR PRINTS; PR00864; PREPILNPTASE.
SQ SEQUENCE 155 AA; 16984 MW; 70BDF2D8543810AC CRC64;

Query Match      87.1%; Score 27; DB 1; Length 155;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
DB 96 WHTWTF 101

RESULT 9
YE63_MYCPN
ID YE63_MYCPN STANDARD; PRT; 157 AA.
AC P75320;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN463 (H08_orf157a).
GN MPN463 OR MF378.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]_TaxID=2104;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).

```

CC -|- SIMILARITY: ALMOST IDENTICAL TO M.PNEUMONIAE MPN091 AND MPN413.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AE000036; AAB96026.1; -;
 DR PIR; S73704; S73704.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 157 AA; 17842 MW; B7F5CCB853B95DB7 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 157;
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 Db 37 WSGWSP 42

RESULT 10

ID YB03 MYCPN STANDARD; PRT; 172 AA.
 AC P75566;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MPN103 (C09_orf172).
 GN MPN103 OR MP051.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelleich R., Hilbert H., Flagens H., Pirkel E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae";
 RL Nucleic Acids Res. 24:4420-4449 (1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000006; AAB95699.1; -;
 DR PIR; S73377; S73377.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 172 AA; 21221 MW; E48E8C70D84PBF9E CRC64;

Query Match 87.1%; Score 27; DB 1; Length 172;
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 Db 37 WTAWFF 42

RESULT 11

VPI_BPP2
 ID VPI_BPP2 STANDARD; PRT; 176 AA.
 AC P26701;
 DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Tail protein I (GPI).
 GN I.
 OS Bacteriophage P2.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC P2-like viruses
 OC NCBI_TaxID=10679;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92165720; PubMed=1531648;
 RA Haggard-Ljungquist E., Halling C., Calendar R.;
 RT "DNA sequences of the tail fiber genes of bacteriophage P2: evidence
 RT for horizontal transfer of tail fiber genes among unrelated
 RT bacteriophages.";
 RL J. Bacteriol. 174:1462-1477 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96036485; PubMed=7483254;
 RA Haggard-Ljungquist E., Jacobsen E., Rishovd S., Six E.W., Nilsen O.,
 RA Sunshine M.G., Lindqvist B.H., Kim K.-J., Barreiro V., Koonin E.V.,
 RA Calendar R.;
 RT "Bacteriophage P2: genes involved in baseplate assembly.";
 RL Virology 213:109-121 (1995).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF063097; AAD03285.1; -;
 DR PIR; A42291; A42291.
 DR InterPro; IPR006521; Tail_P2_I.
 DR TIGRPFAMS; TIGR01634; Tail_P2_I; 1.
 KW Structural protein.
 SQ SEQUENCE 176 AA; 19584 MW; EADA32BCAFCECF1E CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 176;
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 Db 48 WLAWAF 53

RESULT 12
 ARTM_ARTSA
 ID ARTM_ARTSA STANDARD; PRT; 229 AA.
 AC P17720;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Artemin.
 OS Artemia salina (Brine shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 OC Artemiidae; Artemia.
 OX NCBI_TaxID=85549;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91065380; PubMed=2249691;
 RA de Graaf J., Amos R., Moeller W.;
 RT "The primary structure of artemin from Artemia cysts.";
 RL Eur. J. Biochem. 193:737-750 (1990).
 CC -|- DEVELOPMENTAL STAGE: Cyst.
 CC -|- MISCELLANEOUS: ARTEMIN IS ELONGATION FACTOR-ASSOCIATED AND
 CC SELF-ASSOCIATING.
 CC -|- SIMILARITY: Belongs to the ferritin family.
 CC -|- SIMILARITY: Contains 1 ferritin-like diiron domain.
 DR InterPro; IPR008331; Ferritin_Dps.

```

DR InterPro; IPR009040; Ferritin_like.
DR Pfam; PF00210; ferritin; 1.
DR PROSITE; PS00905; FERRITIN LIKE; 1.
KW Iron storage; Iron; Metal-Binding; Acetylation; RNA-binding.
FT DOMAIN 25 173 FERRITIN-LIKE DIIRON.
FT MOD RES 1 1 ACETYLATION.
FT VARIANT 7 7 N -> K.
FT VARIANT 8 8 I -> V.
FT VARIANT 223 223 V -> M.
SQ SEQUENCE 229 AA; 25976 MW; 1A76D149246C25BF CRC64;

Query Match 87.1%; Score 27; DB 1; Length 229;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
Db 50 WSMWAF 55

RESULT 13
YGAZ_ECOLI STANDARD; PRT; 245 AA.
AC P76630;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ygaz.
GN YGAZ OR B2682.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the azlC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000353; AAC75729.1; -.
CC PIR; C65048; C65048.
CC EcoGene; EGI3528; YGAZ.
CC InterPro; IPR004471; AZLC.
CC Pfam; PF03591; AZLC; 1.
KW Hypothetical protein; Transpos; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 64 84 POTENTIAL.
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
SQ SEQUENCE 245 AA; 26107 MW; 22ACE8AB8D7D651B CRC64;

Query Match 87.1%; Score 27; DB 1; Length 245;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
Db 50 WSMWAF 55

RESULT 14
RCEH_RHOCA STANDARD; PRT; 254 AA.
AC P19056;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reaction center protein H chain (Photosynthetic reaction center H
DE subunit).
GN PUHA.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84259352; PubMed=6744416;
RA Youvan D.C., Bylina E.J., Alberti M., Begusch H., Hearst J.E.;
RT "Nucleotide and deduced polypeptide sequences of the photosynthetic
RT reaction-center, B870 antenna, and flanking polypeptides from R.
RT capsulata."
RL Cell 37:949-957 (1984).
CC -!- FUNCTION: THE REACTION CENTER IS A MEMBRANE-BOUND COMPLEX THAT
CC MEDIATES THE INITIAL PHOTOCHEMICAL EVENT IN THE ELECTRON TRANSFER
CC PROCESS OF PHOTOSYNTHESIS.
CC -!- COFACTOR: Binds 4 bacteriochlorophylls, 4 magnesium ions, 2
CC bacteriopheophytins, 2 ubiquinones, and 1 iron ion per trimer.
CC -!- SUBUNIT: Heterotrimer composed of subunits L, M, and H.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K01183; AAA26171.1; -.
CC EMBL; Z11165; CAA7520.1; -.
CC PIR; C28771; C28771.
CC HSSP; P11846; 1YST.
CC InterPro; IPR007903; PRC_barrel.
CC InterPro; IPR005652; PRC.
CC Pfam; PF05239; PRC; 1.
CC TIGRFAMs; TIGR01150; puha; 1.
KW Electron transport; Photosynthesis; Reaction center;
KW Bacteriochlorophyll; Transmembrane.
FT TRANSMEM 12 31
SQ SEQUENCE 254 AA; 28536 MW; D69541F237A19039 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 254;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
Db 18 WSMWAF 23

RESULT 15
YN04_YEAST STANDARD; PRT; 350 AA.
AC P53844;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 40.7 kDa protein in PIK1-POL2 intergenic region.
GN YNL264C OR N0815.

```

CC Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC NCBI_TaxID=4932;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=S288C / FY1679;
 CC MEDLINE=96310631; PubMed=8740425;
 CC Sen-Gupta M., Lyck R., Fleig U., Niedenthal R.K., Hegemann J.H.;
 CC "The sequence of a 24,152 bp segment from the left arm of chromosome
 CC XIV from Saccharomyces cerevisiae between the BNI1 and the POL2
 CC genes.";
 CC Yeast 12:505-514(1996).
 CC RL
 CC -1- SIMILARITY: Contains 1 CRAL-TRIO domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z69381; CAA93367.1; -;
 CC EMBL; Z71507; CAA96136.1; -;
 CC PIR; S63197; S63197.
 CC Germonline; 143237; -;
 CC SGD; S0005175; PDR16.
 CC GO; GO:0005737; Cytoplasm; IDA.
 CC GO; GO:0008526; F:phosphatidylinositol transporter activity; IDA.
 CC GO; GO:0015914; P:phospholipid transport; IDA.
 CC GO; GO:0042493; P:response to drug; IGI.
 CC GO; GO:0016126; P:sterol biosynthesis; IMP.
 CC InterPro; IPR001251; CRAL_TRIO_C.
 CC InterPro; IPR008273; CRAL_TRIO_N.
 CC Pfam; PF00650; CRAL_TRIO; 1.
 CC Pfam; PF03765; CRAL_TRIO_N; 1.
 CC SMART; SM00516; SEC14; 1.
 CC PROSITE; PS0191; CRAL_TRIO; 1.
 CC Hypothetical protein.
 CC KW
 CC DOMAIN 135 295 CRAL-TRIO
 CC FT SEQUENCE 351 AA; 40714 MW; 24C5B3262016F037 CRC64;
 CC SQ
 CC
 CC Query Match 87.1%; Score 27; DB 1; Length 351;
 CC Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 1 WXXWXP 6
 CC Db 247 WLAWTF 252
 CC
 CC RESULT 17
 CC TWHH BRARE STANDARD; PRT; 416 AA.
 CC ID TWHH BRARE
 CC AC Q90419;
 CC DT 15-JUL-1999 (Rel. 38, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Tiggly-winkle hedgehog protein precursor (TWHH).
 CC TWHH.
 CC OS Brachydanio rerio (Zebrafish) (Danio rerio).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 CC OC Cyprinidae; Danio.
 CC OX NCBI_TaxID=7955;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 CC RC TISSUE=Embryo;
 CC RX MEDLINE=96014264; PubMed=7583153;
 CC RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P.,
 CC Porter J.A., Moon R.T., Beachy P.A.;
 CC RT "Patterning activities of vertebrate hedgehog proteins in the
 CC developing eye and brain."
 CC RL Curr. Biol. 5:944-955(1995).
 CC CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL
 CC PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING
 CC EYES
 CC -1- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
 CC cell, while the N-terminal peptide remains associated with the
 CC cell surface. Is also secreted in either cleaved or uncleaved form
 CC to mediate signaling to other cells (By similarity).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
 CC TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN
 CC DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE

OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC NCBI_TaxID=4932;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=S288C / FY1679;
 CC MEDLINE=96310631; PubMed=8740425;
 CC Sen-Gupta M., Lyck R., Fleig U., Niedenthal R.K., Hegemann J.H.;
 CC "The sequence of a 24,152 bp segment from the left arm of chromosome
 CC XIV from Saccharomyces cerevisiae between the BNI1 and the POL2
 CC genes.";
 CC Yeast 12:505-514(1996).
 CC RL
 CC -1- SIMILARITY: Contains 1 CRAL-TRIO domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X2494; CAA63233.1; -;
 CC EMBL; Z71540; CAA96171.1; -;
 CC PIR; S60917; S60917.
 CC Germonline; 143270; -;
 CC SGD; S0005208; PDR17.
 CC GO; GO:0005737; Cytoplasm; IDA.
 CC GO; GO:0008526; F:phosphatidylinositol transporter activity; IDA.
 CC GO; GO:0008654; P:phospholipid biosynthesis; IMP.
 CC GO; GO:0015914; P:phospholipid transport; IDA.
 CC GO; GO:0042493; P:response to drug; IGI.
 CC InterPro; IPR001251; CRAL_TRIO_C.
 CC InterPro; IPR008273; CRAL_TRIO_N.
 CC Pfam; PF00650; CRAL_TRIO; 1.
 CC Pfam; PF03765; CRAL_TRIO_N; 1.
 CC SMART; SM00516; SEC14; 1.
 CC PROSITE; PS0191; CRAL_TRIO; 1.
 CC Hypothetical protein.
 CC KW
 CC DOMAIN 139 297 CRAL-TRIO.
 CC FT SEQUENCE 350 AA; 40678 MW; C18F0D225A69DB86 CRC64;
 CC SQ
 CC
 CC Query Match 87.1%; Score 27; DB 1; Length 350;
 CC Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 1 WXXWXP 6
 CC Db 249 WFAWAF 254
 CC
 CC RESULT 16
 CC YNXL YEAST
 CC ID YNXL YEAST STANDARD; PRT; 351 AA.
 CC AC P53860;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Hypothetical 40.7 kDa protein in CSL4-URE2 intergenic region.
 CC GN YNL231C OR N1158.
 CC OS Saccharomyces cerevisiae (Baker's yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC OX NCBI_TaxID=4932;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=97051596; PubMed=8896273;
 CC RA Pandolfo D., de Antoni A., Lanfranchi G., Valle G.;
 CC RT "The DNA sequence of cosmid 14-5 from chromosome XIV reveals 21 open
 CC RT reading frames including a novel gene encoding a globin-like
 CC RT domain.";
 CC RL Yeast 12:1071-1076(1996).

CC CC DIENCEPHALON. NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.
 CC CC -1- PTM: The C-terminal domain displays an autoprotoecolysis activity
 CC CC and a cholesterol transferase activity. Both activities result in
 CC CC the cleavage of the full-length protein and covalent attachment of
 CC CC a cholesterol moiety to the C-terminal of the newly generated N-
 CC CC terminal fragment (N-product). This covalent modification appears
 CC CC to play an essential role in restricting the spatial distribution
 CC CC of the protein activity to the cell surface. The N-product is the
 CC CC active species in both local and long-range signaling, whereas the
 CC CC C-product has no signaling activity (By similarity).
 CC CC -1- SIMILARITY: Belongs to the hedgehog family.
 CC CC
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC
 CC CC EMBL; U30710; AAC59741.1; -.
 CC CC HSSP; Q62226; 1VHH.
 CC CC MEROPS; C46.001; -.
 CC CC ZFIN; ZDB-GENE-980526-41; twhh.
 CC CC InterPro; IPR009045; Hedgehog/DD pept.
 CC CC InterPro; IPR003587; Hedgehog hint N.
 CC CC InterPro; IPR003586; Hedgehog hint C.
 CC CC InterPro; IPR000320; HH signal.
 CC CC InterPro; IPR001767; Pept C46 hint.
 CC CC InterPro; IPR001657; Peptidase_C46.
 CC CC Pfam; PF01085; HH signal; 1.
 CC CC PRINTS; PR00632; SONICHOOG.
 CC CC ProDom; PD003042; HH_signal; 1.
 CC CC SMART; SM00306; HintN; 1.
 CC CC SWART; SW00306; HintN; 1.
 CC CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC CC Signal; Lipoprotein; Palmitate.
 CC CC FT SIGNAL 1 26
 CC CC FT CHAIN 27 416
 CC CC FT CHAIN 27 200
 CC CC FT CHAIN 201 416
 CC CC FT SITE 200 201
 CC CC FT SITE 270 270
 CC CC INVOLVED IN AUTO-CLEAVAGE (BY
 CC CC SIMILARITY).
 CC CC ESSENTIAL FOR AUTO-CLEAVAGE (BY
 CC CC SIMILARITY).
 CC CC N-palmitoyl cysteine (By similarity).
 CC CC Cholesterol glycine ester (By
 CC CC similarity).
 CC CC SQ SEQUENCE 416 AA; 46576 MW; 61EC2218309CFE59 CRC64;
 CC CC
 CC CC Query March 87.1%; Score 27; DB 1; Length 416;
 CC CC Best Local Similarity 50.0%; Pred. No. 3.3e+02;
 CC CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 CC CC
 CC CC Qy 1 WXXWKF 6
 CC CC Db 354 WAHWAF 359
 CC CC
 CC CC RESULT 18
 CC CC SHH_CHICK STANDARD; PRT; 425 AA.
 CC CC ID SHH_CHICK
 CC CC AC Q91035;
 CC CC DT 15-JUL-1999 (Rel. 38, Created)
 CC CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC CC DE Sonic hedgehog protein precursor (SHH).
 CC CC GN SHH.
 CC CC OS Gallus gallus (Chicken).
 CC CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

CC CC Gallus.
 CC CC NCBI_TaxID=9031;
 CC CC [1]
 CC CC SEQUENCE FROM N.A.
 CC CC TISSUE=Limb bud;
 CC CC MEDLINE=94094333; PubMed=8269518;
 CC CC Riddle R.D., Johnson R.L., Laufer E., Tabin C.;
 CC CC "Sonic hedgehog mediates the polarizing activity of the ZPA.";
 CC CC Cell 75:1401-1416(1993).
 CC CC [2]
 CC CC FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.
 CC CC MEDLINE=95254654; PubMed=7736596;
 CC CC Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
 CC CC Beachy P.A., Jessell T.M.;
 CC CC "Floor plate and motor neuron induction by different concentrations of
 CC CC the amino-terminal cleavage product of sonic hedgehog
 CC CC autoprotoecolysis.";
 CC CC Cell 81:445-455(1995).
 CC CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC CC FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.
 CC CC -1- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
 CC CC cell, while the N-terminal peptide remains associated with the
 CC CC cell surface. Is also secreted in either cleaved or uncleaved form
 CC CC to mediate signaling to other cells (By similarity).
 CC CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD
 CC CC MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE
 CC CC OF THE NEURAL TUBE.
 CC CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE
 CC CC INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE
 CC CC EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF
 CC CC POLARIZING ACTIVITY (ZPA).
 CC CC -1- INDUCTION: By retinoic acid.
 CC CC -1- PTM: The C-terminal domain displays an autoprotoecolysis activity
 CC CC and a cholesterol transferase activity. Both activities result in
 CC CC the cleavage of the full-length protein and covalent attachment of
 CC CC a cholesterol moiety to the C-terminal of the newly generated N-
 CC CC terminal fragment (N-product). This covalent modification appears
 CC CC to play an essential role in restricting the spatial distribution
 CC CC of the protein activity to the cell surface. The N-product is the
 CC CC active species in both local and long-range signaling, whereas the
 CC CC C-product has no signaling activity.
 CC CC -1- SIMILARITY: Belongs to the hedgehog family.
 CC CC
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC
 CC CC EMBL; L28099; AAA72428.1; -.
 CC CC PIR; A49424; A49424.
 CC CC HSSP; Q62226; 1VHH.
 CC CC MEROPS; C46.001; -.
 CC CC InterPro; IPR009045; Hedgehog/DD pept.
 CC CC InterPro; IPR003587; Hedgehog hint N.
 CC CC InterPro; IPR003586; Hedgehog hint C.
 CC CC InterPro; IPR000320; HH signal.
 CC CC InterPro; IPR006141; Intein S.
 CC CC InterPro; IPR001767; Pept_C46 hint.
 CC CC InterPro; IPR001657; Peptidase_C46.
 CC CC Pfam; PF01085; HH_signal; 1.

DR Pfam; PF01079; Hint; 1.
 DR PRINTS; PR00632; SONICHHOG.
 DR ProDom; PD003042; HH signal; 1.
 DR SMART; SM00305; HintC; 1.
 DR SMART; SM00306; HintN; 1.
 DR PROSITE; PS50817; INTEIN N TER; 1.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 26
 FT CHAIN 27 425
 FT CHAIN 27 200
 FT CHAIN 201 425
 FT DOMAIN 390 393
 FT SITE 200 201
 FT SITE 246 246
 FT SITE 270 270
 FT ACT_SITE 273 273
 FT ACT_SITE 27 27
 FT LIPID 200 200
 FT LIPID 200 200
 SQ SEQUENCE 425 AA; 46474 MW; DA9627443DA0173 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 425;
 Best Local Similarity 50.0%; Pred. No. 3.4e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 Db 361 WAHWF 366
 RESULT 19
 SHH_CYNPY STANDARD; PRT; 432 AA.
 ID SHH_CYNPY STANDARD; PRT; 432 AA.
 AC Q90385;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH).
 GN SHH.
 OS Cynops pyrrhogaster (Japanese common newt).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Cynops.
 OX NCBI_TaxID=8330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=96136334; PubMed=8573168;
 RA Takabatake T., Takahashi T.C., Inoue K., Ogawa M., Takeshima K.;
 RT "Activation of two Cynops genes, fork head and sonic hedgehog, in
 animal cap explants.";
 RL Biochem. Biophys. Res. Commun. 218:395-401(1996).
 CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 IN LIMB FORMATION. PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
 TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
 cell, while the N-terminal peptide remains associated with the
 cell surface. Is also secreted in either cleaved or uncleaved form
 to mediate signaling to other cells (By similarity).
 CC -!- INDUCTION: ACTIVATED BY ACTIVIN, BASIC FIBROBLAST GROWTH FACTOR
 (bFGF) AND FORK HEAD.
 CC -!- PM: The C-terminal domain displays an autotranspeptidase activity
 and a cholesterol transferase activity. Both activities result in
 the cleavage of the full-length protein and covalent attachment of

CC a cholesterol moiety to the C-terminal of the newly generated N-
 CC terminal fragment (N-product). This covalent modification appears
 CC to play an essential role in restricting the spatial distribution
 CC of the protein activity to the cell surface. The N-product is the
 CC active species in both local and long-range signaling, whereas the
 CC C-product has no signaling activity (By similarity).
 CC -!- SIMILARITY: Belongs to the hedgehog family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D63339; BAA09657.1; -.
 DR HSSP; O62226; 1VHH.
 DR MEROPS; C46.002; -.
 DR InterPro; IPR009045; Hedgehog/DD_pept.
 DR InterPro; IPR003587; Hedgehog_hint_N.
 DR InterPro; IPR003586; Hedgehog_hintC.
 DR InterPro; IPR00320; HH signal.
 DR InterPro; IPR006141; Intein S.
 DR InterPro; IPR001767; Pept_C46_hint.
 DR InterPro; IPR001657; Peptidase_C46.
 DR Pfam; PF01085; HH_signal; 1.
 DR Pfam; PF01079; Hint; 1.
 DR PRINTS; PR00632; SONICHHOG.
 DR ProDom; PD003042; HH signal; 1.
 DR SMART; SM00305; HintC; 1.
 DR SMART; SM00306; HintN; 1.
 DR PROSITE; PS50817; INTEIN N TER; 1.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 26
 FT CHAIN 27 432
 FT CHAIN 27 200
 FT CHAIN 201 432
 FT SITE 200 201
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT LIPID 27 27
 FT LIPID 200 200
 FT LIPID 200 200
 SQ SEQUENCE 432 AA; 47847 MW; B455C7E746C8E5A8 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 432;
 Best Local Similarity 50.0%; Pred. No. 3.4e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 Db 365 WAHWF 370
 RESULT 20
 TRPE_MYCTU STANDARD; PRT; 516 AA.
 ID TRPE_MYCTU STANDARD; PRT; 516 AA.
 AC Q06127;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Anthranilate synthase component I (EC 4.1.3.27).
 GN TRPE OR RV1609 OR MT1644 OR MYC01B2.01 OR MTV046.07 OR MBL1635.
 OS Mycobacterium tuberculosis, and
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773, 1765;
 RN [1]

```

SEQUENCE FROM N.A.
SPCIES=M.tuberculosis; STRAIN=H37Rv;
MEDLINE=98293987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
[2]
SEQUENCE FROM N.A.
SPCIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
[3]
SEQUENCE FROM N.A.
SPCIES=M.tuberculosis; STRAIN=AF2122/97;
MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
[4]
CATALYTIC ACTIVITY: Chorismate + L-glutamine = anthranilate +
pyruvate + L-glutamate.
[5]
PATHWAY: Tryptophan biosynthesis; first step.
[6]
SUBUNIT: Tetramer of two components I and two components II (By
similarity).
[7]
MISCELLANEOUS: Component I catalyzes the formation of anthranilate
using ammonia rather than glutamine, whereas component II provides
glutamine amidotransferase activity.
[8]
SIMILARITY: Belongs to the anthranilate synthase component I
family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z95554; CAB08903.1; --
EMBL; AE007029; AAK45913.1; --
EMBL; BX248339; CAD96303.1; --
PIR; G70556; G70556.
HSP; P00897; 117Q.
TIGR; MT1644; --.
Tuberculat; RV1609; --.
InterPro; IPR005801; Anth_synth_chor.
InterPro; IPR006805; Anth_synth_I_N.
InterPro; IPR005256; Anth_synth_I.
Pfam; PF04715; Anth_synth_I_N; 1.
Pfam; PF00425; chorismate_Bind; 1.
PRINTS; PR00095; ANTSNTHASEI.
ProDom; PD000779; Anth_synth_chor; 1.
TIGRFAMs; TIGR00564; tpe_mob1; 1.
KW Tryptophan biosynthesis; Lyase; Complete proteome.
SEQUENCE 516 AA; 55848 MW; EF19CDDDE80E802C7 CRC64;

```

```

Query Match 87.1%; Score 27; DB 1; Length 516;
Best Local Similarity 50.0%; Pred. No. 3.9e+02; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 3;
Qy 1 WXXWF 6
Db 63 WSRWSP 68

RESULT 21
TRPE MYCLE
ID TRPE MYCLE STANDARD; PRT; 529 AA.
AC Q9X7C5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anthranilate synthase component I (EC 4.1.3.27).
GN TRPE OR ML1269 OR MLCB1610.31.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
[1]
CATALYTIC ACTIVITY: Chorismate + L-glutamine = anthranilate +
pyruvate + L-glutamate.
[2]
PATHWAY: Tryptophan biosynthesis; first step.
[3]
SUBUNIT: Tetramer of two components I and two components II (By
similarity).
[4]
MISCELLANEOUS: Component I catalyzes the formation of anthranilate
using ammonia rather than glutamine, whereas component II provides
glutamine amidotransferase activity.
[5]
SIMILARITY: Belongs to the anthranilate synthase component I
family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AL049913; CAB43177.1; --
EMBL; AL583921; CAC31650.1; --
PIR; T45254; T45254.
HSP; Q06128; IQDL.
DR Leproma; ML1269; --.
InterPro; IPR005801; Anth_synth_chor.
InterPro; IPR006805; Anth_synth_I_N.
InterPro; IPR005256; Anth_synth_I.
Pfam; PF04715; Anth_synth_I_N; 1.
Pfam; PF00425; chorismate_Bind; 1.
PRINTS; PR00095; ANTSNTHASEI.
ProDom; PD000779; Anth_synth_chor; 1.
TIGRFAMs; TIGR00564; tpe_mob1; 1.
KW Tryptophan biosynthesis; Lyase; Complete proteome.
SEQUENCE 529 AA; 57031 MW; A0D1E0920BA5E3D0 CRC64;

```

Query Match

87.1%; Score 27; DB 1; Length 529;

Best Local Similarity 50.0%; Pred. No. 4e+02; Mismatches 3; Indels 0; Gaps 0;
Matches 3; Conservative 0;

QY 1 WXXWKF 6
| | | |
Db 63 WQWSF 68

RESULT 22

AM01_ARTS1 STANDARD; PRT; 648 AA.
AC Q07121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Copper amine oxidase precursor (EC 1.4.3.6) (MAOXI).
GN MAOI.
OS Arthrobacter sp. (strain P1).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=47915;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 10-54; 358-381 AND 456-466.
RX MEDLINE=93374858; PubMed=8366046;
RA Zhang X., Fuller J.H., McIntire W.S.;
RT "Cloning, sequencing, expression, and regulation of the structural
gene for the copper/topa quinone-containing methylamine oxidase from
Arthrobacter strain P1, a Gram-positive facultative methylotroph.";
RT Arthrobacter strain P1, a Gram-positive facultative methylotroph.";
RL J. Bacteriol. 175:5617-5627(1993).
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
H(2)O(2).
CC -1- COFACTOR: Binds 1 copper ion and 1 topaquinone per subunit.
CC -1- PTM: Topaquinone (TPQ) is generated by copper-dependent
autoxidation of a specific tyrosyl residue (By similarity).
CC -1- SIMILARITY: Belongs to the copper/topaquinone oxidase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L12983; AAA22076.1; -.
CC HSSP; P46881; 1A4V4.
CC InterPro; IPR000269; CuNH oxidase.
CC Pfam; PF01179; Cu_amine_oxid; 1.
CC Pfam; PF02727; Cu_amine_oxidn2; 1.
CC Pfam; PF02728; Cu_amine_oxidn3; 1.
CC PROSITE; PS01164; COPPER AMINE OXID 1; 1.
CC PROSITE; PS01165; COPPER AMINE OXID 2; 1.
KW Oxidoreductase; Copper; TPQ; Metal-Binding.
FT PROPEP 1 9
FT CHAIN 10 648 COPPER AMINE OXIDASE.
FT MOD RES 385 385 TOPAQUINONE (BY SIMILARITY).
FT METAL 436 436 COPPER (POTENTIAL).
FT METAL 438 438 COPPER (POTENTIAL).
FT METAL 595 595 COPPER (POTENTIAL).
SQ SEQUENCE 648 AA; 72760 MW; B2F9E267492253B5 CRC64;
Query Match 87.1%; Score 27; DB 1; Length 648;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWKF 6
| | | |
Db 243 WADWSF 248

RESULT 23

AM02_ARTS1

ID AMO2_ARTS1 STANDARD; PRT; 648 AA.
AC Q07123;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Copper methylamine oxidase precursor (EC 1.4.3.6) (MAOXII).
GN MAOII.
OS Arthrobacter sp. (strain P1).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=47915;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 10-54; 358-381 AND 456-466.
RX MEDLINE=93374858; PubMed=8366046;
RA Zhang X., Fuller J.H., McIntire W.S.;
RT "Cloning, sequencing, expression, and regulation of the structural
gene for the copper/topa quinone-containing methylamine oxidase from
Arthrobacter strain P1, a Gram-positive facultative methylotroph.";
RT Arthrobacter strain P1, a Gram-positive facultative methylotroph.";
RL J. Bacteriol. 175:5617-5627(1993).
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
H(2)O(2).
CC -1- COFACTOR: Binds 1 copper ion and 1 topaquinone per subunit.
CC -1- SUBUNIT: Homodimer.
CC -1- INDUCTION: By methylamine.
CC -1- PTM: Topaquinone (TPQ) is generated by copper-dependent
autoxidation of a specific tyrosyl residue (By similarity).
CC -1- SIMILARITY: Belongs to the copper/topaquinone oxidase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L12990; AAA22074.1; -.
CC HSSP; P46846; A48646.
CC InterPro; IPR000269; CuNH oxidase.
CC Pfam; PF01179; Cu_amine_oxid; 1.
CC Pfam; PF02727; Cu_amine_oxidn2; 1.
CC Pfam; PF02728; Cu_amine_oxidn3; 1.
CC PROSITE; PS01164; COPPER AMINE OXID 1; 1.
CC PROSITE; PS01165; COPPER AMINE OXID 2; 1.
KW Oxidoreductase; Copper; TPQ; Metal-Binding.
FT PROPEP 1 9
FT CHAIN 10 648 COPPER METHYLAMINE OXIDASE.
FT MOD RES 385 385 TOPAQUINONE (BY SIMILARITY).
FT METAL 436 436 COPPER (POTENTIAL).
FT METAL 438 438 COPPER (POTENTIAL).
FT METAL 595 595 COPPER (POTENTIAL).
SQ SEQUENCE 648 AA; 72805 MW; B2FB2787492253B5 CRC64;
Query Match 87.1%; Score 27; DB 1; Length 648;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWKF 6
| | | |
Db 243 WADWSF 248

RESULT 24

ITB8_HUMAN

ID ITB8_HUMAN STANDARD; PRT; 769 AA.
AC P26012;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Integrin beta-8 precursor.
GN ITGB8.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Placenta;
 RX MEDLINE=92011767; PubMed=1918072;
 RA Moyle M., Napier M.A., McLean J.W.;
 RT "Cloning and expression of a divergent integrin subunit beta 8.";
 RL J. Biol. Chem. 266:19650-19658(1991).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP Tin-Wollam A., Sutterer C., Fronick B.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INTEGRIN ALPHA-V/BETA-8 IS A RECEPTOR FOR FIBRONECTIN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-8
 CC ASSOCIATES WITH ALPHA-V.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: PLACENTA, KIDNEY, BRAIN, OVARY, UTERUS, AND IN
 CC SEVERAL TRANSFORMED CELLS. TRANSIENTLY EXPRESSED IN 293 HUMAN
 CC EMBRYONIC KIDNEY CELLS.
 CC -!- SIMILARITY: Belongs to the integrin beta chain family.
 CC -!- SIMILARITY: Contains 2 VWFA-like domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M73780; AAA36034.1; -;
 DR EMBL; AC004130; RAQ96845.1; -;
 DR F01; A41029; A41029.
 DR HSSP; P05106; 1JV2.
 DR Genew; HGNC:6163; ITGB8.
 DR MIM; 604160; -;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0007155; F:cell adhesion; TAS.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR002369; Integrin_B.
 DR InterPro; IPR001169; Integrin_beta_C.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002035; VWFA.
 DR Pfam; PF00362; integrin_B; 1.
 DR PRINTS; PR01186; INTEGRINB.
 DR ProDom; PD001811; Integrin_B; 1.
 DR SMART; SM00187; INB; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00327; VWFA; 1.
 DR PROSITE; PS00243; INTEGRIN_BETA; 2.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Signal.
 FT SIGNAL 1 42
 FT CHAIN 43 769
 FT DOMAIN 43 684
 FT TRANSMEM 685 704
 FT DOMAIN 705 769
 FT DOMAIN 146 384
 FT DOMAIN 471 629
 FT REPEAT 471 510
 FT REPEAT 511 552
 FT REPEAT 553 592
 FT REPEAT 593 629
 FT REPEAT 629 699
 FT DISULFID 47 469
 FT DISULFID 55 65
 FT DISULFID 58 94
 FT DISULFID 68 83
 FT DISULFID 211 218

FT DISULFID 266 307 BY SIMILARITY.
 FT DISULFID 407 419 BY SIMILARITY.
 FT DISULFID 467 471 BY SIMILARITY.
 FT DISULFID 494 499 BY SIMILARITY.
 FT DISULFID 526 531 BY SIMILARITY.
 FT DISULFID 528 561 BY SIMILARITY.
 FT DISULFID 533 546 BY SIMILARITY.
 FT DISULFID 567 572 BY SIMILARITY.
 FT DISULFID 574 583 BY SIMILARITY.
 FT DISULFID 585 593 BY SIMILARITY.
 FT DISULFID 607 612 BY SIMILARITY.
 FT DISULFID 609 657 BY SIMILARITY.
 FT DISULFID 614 624 BY SIMILARITY.
 FT DISULFID 627 630 BY SIMILARITY.
 FT DISULFID 634 643 BY SIMILARITY.
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 431 431 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 769 AA; 85631 MW; F7E3994F92B12A65 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 769;
 Best Local Similarity 50.0%; Pred. No. 5.4e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWVF 6
 DB 28 WAAWVF 33
 RESULT 25
 RAA3 CHLRE
 ID RAA3 CHLRE STANDARD; PRT; 1783 AA.
 AC Q9PEC4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Trans-splicing factor Raa3, chloroplast precursor.
 GN RAA3.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=137c / CC-125;
 RX MEDLINE=21181833; PubMed=11285239;
 RA Rivier C., Goldschmidt-Clermont M., Rochaix J.-D.;
 RT "Identification of an RNA-protein complex involved in chloroplast
 group II intron trans-splicing in Chlamydomonas reinhardtii.";
 RL EMBO J. 20:1765-1773(2001).
 CC -!- FUNCTION: Required for trans-splicing of exons 1 and 2 of the
 CC chloroplast encoded psaa mRNA (a group II intron). May be required
 CC for stability of the chloroplast RNA-protein complex in which it
 CC is found.
 CC -!- SUBUNIT: Part of a 1700 kDa complex that includes the
 CC precursor RNA to exon 1 and the tscA RNA.
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -!- DOMAIN: The N-terminal 453 amino acids are dispensable, while the
 CC C-terminal 630 amino acids are required for function.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF310675; AAG40000.1; -;

```

DR EMBL: AF310674; AAG39999.1; -.
KW Chloroplast; Transit peptide; mRNA processing; mRNA splicing.
FT TRANSIT 1 40
FT CHAIN 41 1783 CHLOROPLAST (POTENTIAL).
FT DOMAIN 112 311 ALA-RICH.
FT DOMAIN 343 436 SER-RICH.
FT DOMAIN 478 1078 ALA-RICH.
FT DOMAIN 1310 1415 ALA-RICH.
FT DOMAIN 1416 1430 GLN-RICH.
FT DOMAIN 1496 1506 ARG-RICH.
FT DOMAIN 1678 1722 ALA-RICH.
FT DOMAIN 771 778 POLY-ALA.
FT DOMAIN 920 926 POLY-PRO.
FT DOMAIN 927 932 POLY-ALA.
FT DOMAIN 1047 1055 POLY-GLY.
FT DOMAIN 1318 1325 POLY-ALA.
FT DOMAIN 1405 1415 POLY-ALA.
FT DOMAIN 1669 1675 POLY-PRO.
SQ SEQUENCE 1783 AA; 180399 MW; 40F6206BA6EBDCDB CRC64;

Query Match 87.1%; Score 27; DB 1; Length 1783;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
| | |
Db 1010 WSLWAF 1015

RESULT 26
YKFF_ECOLI STANDARD; PRT; 79 AA.
AC P75677; Q9R2D8;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ykff.
GS YKFF OR B0249.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO E.COLI YPJ1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000133; AAC73352.1; -.
CC EMBL: D83536; BAA77918.1; ALT INIT.

```

```

DR PIR: A64750; A64750.
DR Ecogene; EGI4283; Ykff.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 79 AA; 9014 MW; 614D5749721A234A CRC64;

Query Match 83.9%; Score 26; DB 1; Length 79;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
| | |
Db 50 WEANWF 55

RESULT 27
YPJ1_ECOLI STANDARD; PRT; 90 AA.
AC P58095;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ypj1.
GN YPJ1 OR B2641.1.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP IDENTIFICATION, AND CONCEPTUAL TRANSLATION.
RA Rudd K.E.;
RL Unpublished observations (JAN-1999).
CC -1- SIMILARITY: STRONG, TO E.COLI YKFF.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION. A FRAMESHIFT HAD TO BE
CC CORRECTED IN POSITION 60 AND TWO IN FRAME STOP CODONS HAD TO BE
CC SUPPRESSED IN POSITIONS 8 AND 11.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000349; -. NOT ANNOTATED_CDS.
CC DR Ecogene; EGI3309; YPJ1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 10070 MW; E1C3E2F6C2E8B697 CRC64;

Query Match 83.9%; Score 26; DB 1; Length 90;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
| | |
Db 61 WRAWNF 66

RESULT 28
YG50_MYCPN STANDARD; PRT; 101 AA.
AC P75147;
DT 16-OCT-2001 (Rel. 40, Created)

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical lipoprotein MPN650 precursor (E09_orf101).
 GN MPN650 OR MP192.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Hammelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Potential).
 CC -!- SIMILARITY: Belongs to the MG439 / MG440 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB000020; AAB95840.1; -.
 DR PIR; S73518; S73518.
 DR InterPro; IPR001595; Lipoprotein 3.
 DR InterPro; IPR000437; Prok lipoprot_3.
 DR Pfam; PF00938; Lipoprotein 3; 1.
 DR ProDom; PD003276; Lipoprotein_3; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Lipoprotein; Membrane; Signal;
 KW Complete proteome; Palmitate.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 101 HYPOTHETICAL LIPOPROTEIN MPN650.
 FT LIPID 20 20 N-palmitoyl cysteine (Potential).
 FT LIPID 20 20 S-diacylglycerol cysteine (Potential).
 FT SEQUENCE 101 AA; 11183 MW; C71F0B9EF197573D CRC64;
 Query Match 83.9%; Score 26; DB 1; Length 101;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 WXXWKF 6
 Db 95 WTNWKF 100
 RESULT 29
 HV05_CARAU ID HV05_CARAU STANDARD; PRT; 116 AA.
 AC P19181;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, last annotation update)
 DE 19 heavy chain v region 5A precursor.
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88144476; PubMed=3125551;
 RA Wilson M.R., Middleton D., Warr G.W.;
 RT "Immunoglobulin heavy chain variable region gene evolution: structure
 RT and family relationships of two genes and a pseudogene in a teleost
 RT fish".
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).

DR HSSP; P01772; 2FB4.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 KW Immunoglobulin V region; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 116 IG HEAVY CHAIN V REGION 5A.
 FT DOMAIN 20 49 FRAMEWORK-1.
 FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 55 68 FRAMEWORK-2.
 FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 85 116 FRAMEWORK-3.
 FT DISULFID 41 114 BY SIMILARITY.
 FT NON_TER 116 116
 SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;
 Query Match 83.9%; Score 26; DB 1; Length 116;
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 WXXWKF 6
 Db 4 WLSWVF 9
 RESULT 30
 CRCB_FALSO ID CRCB_FALSO STANDARD; PRT; 126 AA.
 AC Q8XZR2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein crcb homolog.
 GN CRCB OR RSC1333 OR RS02855.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMI1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brattier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the crcb family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL646064; CAD15035.1; -.
 DR HAMAP; MF_00454; -; 1.
 DR InterPro; IPR003691; Camphor_Crcb.
 DR Pfam; PF02537; CRCB; 1.
 DR TIGRFAMs; TIGR00494; crcb; 1.
 KW Transmembrane; Complete proteome.
 FT TRANSMEM 5 24 POTENTIAL.
 FT TRANSMEM 34 56 POTENTIAL.
 FT TRANSMEM 68 90 POTENTIAL.
 FT TRANSMEM 100 122 POTENTIAL.

```

SQ SEQUENCE 126 AA; 13101 MW; C2443FBAE5C81CB3 CRC64;
Query Match 83.9%; Score 26; DB 1; Length 126;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 18 WLRWAF 23

RESULT 31
MERT_STAAU
ID MERT_STAAU STANDARD; PRT; 128 AA.
AC P06656;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mercuric transport protein (Mercury ion transport protein).
GN MERT.

OS Staphylococcus aureus.
OG Plasmid p1258.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
ON NCBI_TaxID=1280;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87260937; PubMed=3037534;
RA Laddaga R.A., Chu L., Misra T.K., Silver S.;
RT "Nucleotide sequence and expression of the mercurial-resistance
RL operon from Staphylococcus aureus plasmid p1258.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:5106-5110(1987).
CC -!- FUNCTION: INVOLVED IN MERCURIC TRANSPORT. PASSES A HG(2+) ION
CC FROM THE PERIPLASMIC MERTP PROTEIN TO THE MERCURIC REDUCTASE
CC (MERT).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L29436; AAA98244.1; -.
DR PIR; D29504; D29504.
KW Transport; Mercuric resistance; Transmembrane; Mercury; Plasmid.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT METAL 47 47 MERCURY (POTENTIAL).
FT METAL 48 48 MERCURY (POTENTIAL).
FT METAL 114 114 MERCURY (POTENTIAL).
FT METAL 115 115 MERCURY (POTENTIAL).
SQ SEQUENCE 128 AA; 14182 MW; C39974E2B4A8AFE91 CRC64;

Query Match 83.9%; Score 26; DB 1; Length 128;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 28 WGVWAF 33

RESULT 32
Y565_METJA
ID Y565_METJA STANDARD; PRT; 147 AA.
AC Q57985;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0565.
GN MJ0565.

OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
ON NCBI_TaxID=2190;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RA MEDLINE=96337999; PubMed=868087; Zhou L., Fleischmann R.D.,
RA Bult C.J., White O., Olsen G.J., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Adams M.D., Reich C.I.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Merrick J.M., Glodek A.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Fuhmann J.L., Nguyen D.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Sadow P.W., Hanna M.C.,
RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Borodovsky M.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67505; AAB98561.1; -.
DR PIR; E64370; E64370.
DR TIGR; MJ0565; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 45 65 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 116 136 POTENTIAL.
SQ SEQUENCE 147 AA; 15900 MW; A19D2A29BDFB36AB CRC64;

Query Match 83.9%; Score 26; DB 1; Length 147;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 113 WADWGF 118

RESULT 33
YVGO_BACSU
ID YVGO_BACSU STANDARD; PRT; 161 AA.
AC O32211;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB Stress response protein yvgo precursor.
GN YVGO OR BSU33410.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ON NCBI_TaxID=1423;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RA MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

```



```

RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata B., Kasahara Y., Klair-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapides A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Srokun A., Tacconi B., Takagi T., Takahashi H., Takemaru K.,
RA Toste V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256 (1997).
RN [2]
RN TRANSCRIPTIONAL REGULATION.
RP MEDLINE=21984451; PubMed=11988534;
RX Pragai Z., Harwood C.R.;
RA "Regulatory interactions between the pho and sigma(B)-dependent
RT general stress regulons of Bacillus subtilis."
RL Microbiology 148:1593-1602(2002).
CC -!- INDUCTION: By phosphate starvation, via the alternative sigma
CC factor sigma-B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 299121; CAB15346.1; -.
DR PIR; D70040; D70040.
DR Subtilist; BG14097; yvgo.
DR InterPro; IPR008972; Cupredoxin.
KW SIGNAL; Complete proteome.
FT CHAIN 27 161 STRESS RESPONSE PROTEIN YVGO.
FT SEQUENCE 161 AA; 17602 MW; 76729851E5942307 CRC64;
QY Query Match 83.9%; Score 26; DB 1; Length 161;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DB 140 WINWAF 145
QY 1 WXXWXP 6
DB 140 WINWAF 145
RESULT 34
Y428_MYCGE STANDARD; PRT; 171 AA.
AC F47667;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG428.
GN MG428.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;

```

```

RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.B., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:397-403(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39725; AAC72449.1; -.
DR PIR; C64247; C64247.
DR TIGR; MG428; -.
DR InterPro; IPR000792; HTH LuxR.
DR InterPro; IPR009043; RNA_pol_sigma.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 171 AA; 20259 MW; 87FB346C334E5E72 CRC64;
QY Query Match 83.9%; Score 26; DB 1; Length 171;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DB 19 WKSNEP 24
QY 1 WXXWXP 6
DB 19 WKSNEP 24
RESULT 35
VG37_BPMU STANDARD; PRT; 182 AA.
AC Q9T1V8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Protein gp37.
GN gp37.
OS Bacteriophage Mu.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC Mu-like viruses.
OX NCBI_TaxID=10677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920971; PubMed=11922669;
RA Morgan G.J., Hatfull G.F., Casjens S., Hendrix R.W.;
RT "Bacteriophage Mu genome sequence: analysis and comparison with
RT Mu-like prophages in Haemophilus, Neisseria and Deinococcus."
RL J. Mol. Biol. 317:337-359(2002).
CC -!- SIMILARITY: STRONG, TO H.INFLUENZA H1509.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF083977; AAF01115.1; -.
DR PIR; 20556 MW; 0B9FC3E226996E20 CRC64;
SQ SEQUENCE 182 AA; 20556 MW;
QY Query Match 83.9%; Score 26; DB 1; Length 182;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWXP 6

```



```

OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA MEDLINE=97426617; PubMed=9278503;
RX Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SIMILARITY: Belongs to the CDP-alcohol phosphatidyltransferase
CC class-I family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL; AE000270; AAC74828.1; ALT_INIT.
CC DR EcoGene; EG14008; ynfJ.
CC DR InterPro; IPR000462; CDP-OH P trans.
CC DR Pfam; PF01066; CDP-OH P trans; 1.
CC DR PROSITE; PS00379; CDP_ALCOHOL_P_TRANSF; FALSE_NEG.
CC DR Hypothetical protein; Transferase; Complete proteome.
CC SQ SEQUENCE 206 AA; 22752 MW; E4C20A0810F90641 CRC64;
CC
CC Query Match 83.9%; Score 26; DB 1; Length 206;
CC Best Local Similarity 50.0%; Pred. No. 2.8e+02;
CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC
CC Qy 1 WXXWKF 6
CC Db 176 WFAWIF 181
CC
CC RESULT 39
CC IF4E_XENLA STANDARD; PRT; 213 AA.
CC AC P48597;
CC DT 01-FEB-1996 (Rel. 33; Created)
CC DT 01-FEB-1996 (Rel. 33; Last sequence update)
CC DT 10-OCT-2003 (Rel. 42; Last annotation update)
CC DE Eukaryotic translation initiation factor 4E (eIF4E) (eIF-4E) (mRNA
CC cap-binding protein) (eIF-4F 25 kDa subunit).
CC OS Xenopus laevis (African clawed frog).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
CC OC Xenopodinae; Xenopus.
CC OX NCBI_TaxID=8355;
CC RN [1]
CC RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
CC RX MEDLINE=95180446; PubMed=7875328;
CC RA Wakiyama M., Saigo M., Shiohara K., Miura K.I.;
CC RT "mRNA encoding the translation initiation factor eIF-4E is expressed
CC early in Xenopus embryogenesis.";
CC RL FEBS Lett. 360:191-193(1995).
CC
CC [2]
CC RP FUNCTION, INTERACTION WITH MASKIN, SUBCELLULAR LOCATION, AND
CC RP DEVELOPMENTAL STAGE.
CC RX MEDLINE=20101243; PubMed=10635326;
CC RA Stebbins-Boaz B., Cao Q., de Moor C.H., Mendez R., Richter J.D.;
CC RT "Maskin is a CPEP-associated factor that transiently interacts with
CC eIF-4E.";
CC RL Mol. Cell 4:1017-1027(1999).
CC
CC -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC mRNA cap during an early step in the initiation of protein
CC synthesis and facilitates ribosome binding by inducing the
CC unwinding of the mRNAs secondary structures. Maternal RNA in
CC oocytes remain in a dormant state as Maskin outcompetes eIF4G to

```

```

CC bind eIF4E, thereby preventing translation. During oocyte
CC maturation this complex dissolves and eIF4G binds eIF4E to allow
CC translation of maternal RNAs.
CC
CC -!- SUBUNIT: eIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least eIF4A, eIF4E and eIF4G. eIF4E is also
CC known to interact with other partners. Maskin competes with eIF4G
CC for binding to eIF4E.
CC
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; when interacting with Maskin.
CC
CC -!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.
CC
CC -!- PTM: Phosphorylation increases the ability of the protein to bind
CC to mRNA caps and to form the eIF4F complex (By similarity).
CC
CC -!- SIMILARITY: Belongs to the eukaryotic initiation factor 4E family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL; D31837; BAA06623.1; -.
CC DR PIR; I51413; I51413.
CC DR HSSP; P07260; IAP8.
CC DR InterPro; IPR001040; TIF_eIF_4E.
CC DR Pfam; PF01652; IF4E; 1.
CC DR ProDom; PD003697; TIF_eIF_4E; 1.
CC DR PROSITE; PS00813; IF4E; 1.
CC DR Protein biosynthesis; Translation regulation; Initiation factor;
CC RNA-binding; Phosphorylation.
CC FT MOD_RES 205 205 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
CC SQ SEQUENCE 213 AA; 24635 MW; E9412DF235AE9E8 CRC64;
CC
CC Query Match 83.9%; Score 26; DB 1; Length 213;
CC Best Local Similarity 50.0%; Pred. No. 2.9e+02;
CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC
CC Qy 1 WXXWKF 6
CC Db 39 WALWFF 44
CC
CC RESULT 40
CC IF4E_APLCA STANDARD; PRT; 215 AA.
CC ID IF4E_APLCA STANDARD; PRT; 215 AA.
CC AC 077210;
CC DT 30-MAY-2000 (Rel. 39; Created)
CC DT 30-MAY-2000 (Rel. 39; Last sequence update)
CC DT 10-OCT-2003 (Rel. 42; Last annotation update)
CC DE Eukaryotic translation initiation factor 4E (eIF4E) (eIF-4E) (mRNA
CC cap-binding protein) (eIF-4F 25 kDa subunit).
CC OS Aplysia californica (California sea hare).
CC OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
CC OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidia;
CC OC Aplysioidea; Aplysiidae; Aplysia.
CC OX NCBI_TaxID=6500;
CC RN [1]
CC RP SEQUENCE FROM N.A., AND PHOSPHORYLATION.
CC RC TISSUE=Nerve;
CC RX MEDLINE=99009053; PubMed=9792652;
CC RA Dyer J.R., Pepio A.M., Yanow S.K., Sossin W.S.;
CC RT "Phosphorylation of eIF4E at a conserved serine in Aplysia.";
CC RL J. Biol. Chem. 273:29469-29474(1998).
CC
CC -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC mRNA cap during an early step in the initiation of protein
CC synthesis and facilitates ribosome binding by inducing the
CC unwinding of the mRNAs secondary structures.
CC
CC -!- SUBUNIT: eIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least eIF4A, eIF4E and eIF4G. eIF4E is also
CC known to interact with other partners (By similarity).
CC
CC -!- PTM: Phosphorylation increases the ability of the protein to bind

```

CC to mRNA caps and to form the EIF4F complex.
 CC -1- SIMILARITY: Belongs to the eukaryotic initiation factor 4E family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF085810; AAC36720.1; -;
 CC HSSP; P07260; IAP8.
 CC InterPro; IPR001040; TIF_eIF_4E.
 CC Pfam; PF01652; IF4E; 1.
 CC ProDom; PD003697; TIF_eIF_4E; 1.
 CC PROSITE; PS00813; IF4E; 1.
 CC Protein biosynthesis; Translation regulation; Initiation factor;
 CC RNA-binding; Phosphorylation.
 CC MOD RES 207 207 PHOSPHORYLATION (BY PKC).
 CC FT RNA-binding; Phosphorylation.
 CC SQ SEQUENCE 215 AA; 24646 MW; 9EE00CB6DE8162E7 CRC64;
 CC
 CC Query Match 83.9%; Score 26; DB 1; Length 215;
 CC Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 1 WXXWKF 6
 CC | | | | |
 CC Db 41 WAWWFF 46
 CC
 CC RESULT 41
 CC IF4E BOVIN
 CC ID IF4E BOVIN STANDARD; PRT; 217 AA.
 CC AC Q2N075;
 CC DT 10-OCT-2003 (Rel. 42, Created)
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Eukaryotic translation initiation factor 4E (eIF4E) (mRNA
 CC DE cap-binding protein).
 CC GN EIF4E.
 CC OS Bos taurus (Bovine).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC OC Bovidae; Bovinae; Bos.
 CC OX NCBI_TaxID=9913;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=21910451; PubMed=11913777;
 CC RA Long E., Capuco A.V., Zhao X.;
 CC RT "Cloning of bovine eukaryotic translation initiation factor 4E
 CC RT (eIF-4E) and its expression in the bovine mammary gland at different
 CC RT physiological stages.";
 CC RL DNA Seq. 12:319-329(2001).
 CC -1- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
 CC mRNA cap during an early step in the initiation of protein
 CC synthesis and facilitates ribosome binding by inducing the
 CC unwinding of the mRNAs secondary structures (By similarity).
 CC -1- SUBUNIT: EIF4F is a multi-subunit complex, the composition of
 CC which varies with external and internal environmental conditions.
 CC It is composed of at least EIF4A, EIF4E and EIF4G. EIF4E is also
 CC known to interact with other partners. The interaction with
 CC EIF4ENIF1 mediates the import into the nucleus (By similarity).
 CC -1- PTM: Phosphorylation increases the ability of the protein to bind
 CC to mRNA caps and to form the EIF4F complex (By similarity).
 CC -1- SIMILARITY: Belongs to the eukaryotic initiation factor 4E family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF257235; AAF66991.1; -;
 CC HSSP; P07260; IAP8.
 CC InterPro; IPR001040; TIF_eIF_4E.
 CC Pfam; PF01652; IF4E; 1.
 CC ProDom; PD003697; TIF_eIF_4E; 1.
 CC PROSITE; PS00813; IF4E; 1.
 CC Protein biosynthesis; Translation regulation; Initiation factor;
 CC RNA-binding; Phosphorylation; Multigene family.
 CC FT MOD RES 209 209 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
 CC SQ SEQUENCE 217 AA; 25063 MW; D2C7D80D9E4815C CRC64;
 CC
 CC Query Match 83.9%; Score 26; DB 1; Length 217;
 CC Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 1 WXXWKF 6
 CC | | | | |
 CC Db 43 WAWWFF 48
 CC
 CC RESULT 42
 CC IF4E HUMAN
 CC ID IF4E HUMAN STANDARD; PRT; 217 AA.
 CC AC P06730; Q96E95;
 CC DT 01-JAN-1988 (Rel. 06, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Eukaryotic translation initiation factor 4E (eIF4E) (mRNA
 CC DE cap-binding protein) (eIF-4F 25 kDa subunit).
 CC GN EIF4E.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Placenta;
 CC RX MEDLINE=87147214; PubMed=3469651;
 CC RA Rychlik W., Domier L.L., Gardner P.R., Hellmann G.M., Rhoads R.E.;
 CC RT "Amino acid sequence of the mRNA cap-binding protein from human
 CC RT tissues.";
 CC RL Proc. Natl. Acad. Sci. U.S.A. 84:945-949(1987).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Brain;
 CC RX MEDLINE=22388257; PubMed=12477932;
 CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 CC RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 CC RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 CC RA Hopkins R.F., Moore H., Moore T., Max S.I., Wang J., Hsieh P.,
 CC RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 CC RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 CC RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 CC RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 CC RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 CC RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 CC RA Vallaloo D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 CC RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 CC RA Whiting M., Madan A., Young A.C., Shewchenko Y., Bouffard G.G.,
 CC RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 CC RA Rodriguez A.C., Guichard J., Schmutz J., Myers R.M.,
 CC RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smaluk D.E.,
 CC RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 CC RT "Generation and initial analysis of more than 15,000 full-length
 CC RT human and mouse cDNA sequences.";
 CC RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC RN [3]
 CC RP PARTIAL SEQUENCE.
 CC RX MEDLINE=91131548; PubMed=1993647;
 CC RA Marino M.M., Feld L.J., Jaffe E.A., Pfeffer L.M., Han Y.-M.,
 CC RA Donner D.B.;

CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF257235; AAF66991.1; -;
 CC HSSP; P07260; IAP8.
 CC InterPro; IPR001040; TIF_eIF_4E.
 CC Pfam; PF01652; IF4E; 1.
 CC ProDom; PD003697; TIF_eIF_4E; 1.
 CC PROSITE; PS00813; IF4E; 1.
 CC Protein biosynthesis; Translation regulation; Initiation factor;
 CC RNA-binding; Phosphorylation; Multigene family.
 CC FT MOD RES 209 209 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
 CC SQ SEQUENCE 217 AA; 25063 MW; D2C7D80D9E4815C CRC64;
 CC
 CC Query Match 83.9%; Score 26; DB 1; Length 217;
 CC Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 1 WXXWKF 6
 CC | | | | |
 CC Db 43 WAWWFF 48
 CC
 CC RESULT 42
 CC IF4E HUMAN
 CC ID IF4E HUMAN STANDARD; PRT; 217 AA.
 CC AC P06730; Q96E95;
 CC DT 01-JAN-1988 (Rel. 06, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Eukaryotic translation initiation factor 4E (eIF4E) (eIF-4E) (mRNA
 CC DE cap-binding protein) (eIF-4F 25 kDa subunit).
 CC GN EIF4E.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Placenta;
 CC RX MEDLINE=87147214; PubMed=3469651;
 CC RA Rychlik W., Domier L.L., Gardner P.R., Hellmann G.M., Rhoads R.E.;
 CC RT "Amino acid sequence of the mRNA cap-binding protein from human
 CC RT tissues.";
 CC RL Proc. Natl. Acad. Sci. U.S.A. 84:945-949(1987).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Brain;
 CC RX MEDLINE=22388257; PubMed=12477932;
 CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 CC RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 CC RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 CC RA Hopkins R.F., Moore H., Moore T., Max S.I., Wang J., Hsieh P.,
 CC RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 CC RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 CC RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 CC RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 CC RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 CC RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 CC RA Vallaloo D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 CC RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 CC RA Whiting M., Madan A., Young A.C., Shewchenko Y., Bouffard G.G.,
 CC RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 CC RA Rodriguez A.C., Guichard J., Schmutz J., Myers R.M.,
 CC RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smaluk D.E.,
 CC RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 CC RT "Generation and initial analysis of more than 15,000 full-length
 CC RT human and mouse cDNA sequences.";
 CC RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC RN [3]
 CC RP PARTIAL SEQUENCE.
 CC RX MEDLINE=91131548; PubMed=1993647;
 CC RA Marino M.M., Feld L.J., Jaffe E.A., Pfeffer L.M., Han Y.-M.,
 CC RA Donner D.B.;

"Phosphorylation of the proto-oncogene product eukaryotic initiation factor 4E is a common cellular response to tumor necrosis factor.";
 J. Biol. Chem. 266:2685-2688(1991).
 [4]
 MUTAGENESIS OF TRP-102; GLU-103; ASP-104 AND GLU-105.
 Ueda H., Iyo H., Doi M., Inoue M., Ishida T., Morioka H., Tanaka T., Nishikawa S., Usugi S.;
 RA MEDLINE=91192132; PubMed=1672854;
 RA "Combination of Trp and Glu residues for recognition of mRNA cap structure. Analysis of m7G base recognition site of human cap binding protein (IF-4E) by site-directed mutagenesis.";
 RA FEBS Lett. 280:207-210(1991).
 [5]
 PHOSPHORYLATION SITE SER-53.
 RX MEDLINE=87280093; PubMed=3112145;
 RA Rychlik W., Russ M.A., Rhoads R.E.;
 RA "Phosphorylation site of eukaryotic initiation factor 4E.";
 RT J. Biol. Chem. 262:10434-10437(1987).
 [6]
 PHOSPHORYLATION SITE SER-53.
 RX MEDLINE=93280157; PubMed=8505316;
 RA Kaufman R.J., Murtha-Riel P., Pittman D.D., Davies M.V.;
 RA "Characterization of wild-type and Ser53 mutant eukaryotic initiation factor 4E overexpression in mammalian cells.";
 RT J. Biol. Chem. 268:11902-11909(1993).
 [7]
 PHOSPHORYLATION SITE SER-53.
 RX MEDLINE=96011649; PubMed=7590282;
 RA Zhang Y., Klein H.L., Schneider R.J.;
 RA "Role of Ser-53 phosphorylation in the activity of human translation initiation factor eIF-4E in mammalian and yeast cells.";
 RL Gene 163:283-288(1995).
 [8]
 PHOSPHORYLATION SITE SER-209.
 RX MEDLINE=95301551; PubMed=7782323;
 RA Joshi B., Cai A.L., Keiper B.D., Minich W.B., Mendez R., Beach C.M., Stepinski J., Stolarski R., Darzykiewicz E., Rhoads R.E.;
 RA "Phosphorylation of eukaryotic protein synthesis initiation factor 4E at Ser-209.";
 RT J. Biol. Chem. 270:14597-14603(1995).
 [9]
 PHOSPHORYLATION SITE SER-209.
 RX MEDLINE=95319427; PubMed=7665584;
 RA Flynn A., Proud C.G.;
 RA "Serine 209, not serine 53, is the major site of phosphorylation in initiation factor eIF-4E in serum-treated Chinese hamster ovary cells.";
 RT J. Biol. Chem. 270:21684-21688(1995).
 [10]
 INTERACTION WITH EIF4G AND EIF4EBP1.
 RX MEDLINE=96091142; PubMed=8521827;
 RA Haghighat A., Mader S., Pause A., Sonenberg N.;
 RA "Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E.";
 RT EMBO J. 14:5701-5709(1995).
 [11]
 INTERACTION WITH EIF4ENIF1.
 RX TISSUE=Fetal brain, and Placenta;
 RX MEDLINE=203115895; PubMed=10856257;
 RA Dostie J., Ferraiuolo M., Pause A., Adam S.A., Sonenberg N.;
 RA "A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E.";
 RT EMBO J. 19:3142-3156(2000).
 [12]
 X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=21868781; PubMed=11879179;
 RA Tomoko K., Shen X., Okabe K., Nozoe Y., Fukuhara S., Morino S., Ishida T., Taniguchi T., Hasegawa H., Terashima A., Sasaki M., Katsuya Y., Kitamura K., Miyoshi H., Ishikawa M., Miura K.;
 RA "Crystal structures of 7-methylguanosine 5'-triphosphate (m(7)GTP) and P(1)-7-methylguanosine-P(3)-adenosine-5',5'-triphosphate (m(7)GppA)-bound human full-length eukaryotic initiation factor 4E:

biological importance of the C-terminal flexible region.";
 Biochem. J. 362:533-544(2002).
 -1- FUNCTION: Recognizes and binds the 7-methylguanosine-containing mRNA cap during an early step in the initiation of protein synthesis and facilitates ribosome binding by inducing the unwinding of the mRNA secondary structures.
 -1- SUBUNIT: EIF4F is a multi-subunit complex, the composition of which varies with external and internal environmental conditions. It is composed of at least EIF4A, EIF4E and EIF4G. EIF4E is also known to interact with other partners. The interaction with EIF4ENIF1 mediates the import into the nucleus. Nonphosphorylated EIF4BBP1, EIF4BBP2 and EIF4BBP3 compete with EIF4G to interact with EIF4E; insulin stimulated MAP-kinase (MAPK1 and MAPK3) phosphorylation of EIF4EBP1 causes dissociation of the complex allowing EIF4G to bind and consequent initiation of translation. Rapamycin can attenuate insulin stimulation, mediated by FXBP3.
 -1- PTM: Phosphorylation increases the ability of the protein to bind to mRNA caps and to form the EIF4F complex.
 -1- SIMILARITY: Belongs to the eukaryotic initiation factor 4E family.
 -1- CAUTION: Was originally thought to be phosphorylated on Ser-53; this was latter shown to be wrong.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

 EMBL; M15353; AAC13647.1; --
 EMBL; BC012611; AAH12611.1; --
 PIR; A26411; A26411.
 PDB; 1IPB; 08-MAY-02.
 PDB; 1IPC; 08-MAY-02.
 Genew; HGNC:3287; EIF4E.
 GK; P06730; --
 MIM; 133440; --
 DR GO; GO:0008304; C:eukaryotic translation initiation factor 4 . . . ; TAS.
 DR GO; GO:0000339; F:RNA cap binding; TAS.
 DR GO; GO:0003743; F:translation initiation factor activity; TAS.
 DR GO; GO:0006441; P:binding to mRNA cap; TAS.
 DR InterPro; IPR01040; TIF_eIF_4E.
 DR Pfam; PF01652; IF4E; 1.
 DR ProDom; PD003697; TIF_eIF_4E; 1.
 DR PROSITE; PS00813; IF4E; 1.
 KW Protein biosynthesis; Translation regulation; Initiation factor;
 KW RNA-binding; Phosphorylation; Multigene family; 3D-structure.
 FT MOD_RES 209 209 PHOSPHORYLATION (BY PKC).
 FT CONFLICT 127 127 D -> N (IN REF. 2).
 SQ SEQUENCE 217 AA; 25097 MW; B869B8DE615B699D CRC64;
 Query Match 83.9%; Score 26; DB 1; Length 217;
 Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 DB 43 WALWFF 48

 RESULT 43
 IF4E_MOUSE
 ID IF4E_MOUSE STANDARD; PRT; 217 AA.
 AC P20415;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Eukaryotic translation initiation factor 4E (eIF4E) (eIF-4E) (mRNA cap-binding protein) (eIF-4F 25 kDa subunit).
 DE Eukaryotic translation initiation factor 4E (eIF4E) (eIF-4E) (mRNA cap-binding protein) (eIF-4F 25 kDa subunit).
 GN EIF4E.
 OS Mus musculus (Mouse), and
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090, 10116;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX MEDLINE=89308629; PubMed=2663851;
RA Altmann M., Mueller P.P., Pelletier J., Sonenberg N., Trachsel H.;
RT "A mammalian translation initiation factor can substitute for its
RT yeast homologue in vivo.";
RL J. Biol. Chem. 264:12145-12147(1989).
RN [2]
RN SEQUENCE FROM N.A.
RP SPECIES=Mouse;
RX MEDLINE=91244820; PubMed=2037592;
RA Jaramillo M., Pelletier J., Edery I., Nielsen P.J., Sonenberg N.;
RT "Multiple mRNAs encode the murine translation initiation factor
RT eIF-4E.";
RL J. Biol. Chem. 266:10446-10451(1991).
RN [3]
RN SEQUENCE FROM N.A.
RP SPECIES=Mouse; STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE FROM N.A., FUNCTION, PHOSPHORYLATION, AND TISSUE SPECIFICITY.
RC SPECIES=Rat; TISSUE=Testis;
RX MEDLINE=96109128; PubMed=8558852;
RA Miyagi Y., Kerr S., Sugiyama A., Asai A., Shibuya M., Fujimoto H.,
RA Kuchino Y.;
RT "Abundant expression of translation initiation factor eIF-4E in post-
RT meiotic germ cells of the rat testis.";
RL Lab. Invest. 73:890-898(1995).
RN [5]
RN FUNCTION, AND INTERACTION WITH EIF4BP1.
RC SPECIES=Rat;
RX MEDLINE=95025978; PubMed=7939721;
RA Lin T.-A., Kong X., Heystead T.A.J., Pause A., Belsham G.J.,
RA Sonenberg N., Lawrence J.C. Jr.;
RT "PHAS-I as a link between mitogen-activated protein kinase and
RT translation initiation.";
RL Science 266:653-656(1994).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 28-217.
RC SPECIES=Mouse;
RX MEDLINE=99322663; PubMed=10394359;
RA Marcotrigiano J., Gingras A.C., Sonenberg N., Burley S.K.;
RT "Cap-dependent translation initiation in eukaryotes is regulated by a
RT molecular mimic of eIF4G.";
RL Mol. Cell 3:707-716(1999).
CC -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC mRNA cap during an early step in the initiation of protein
CC synthesis and facilitates ribosome binding by inducing the
CC unwinding of the mRNAs secondary structures. May play an important

CC role in spermatogenesis through translational regulation of stage-
CC specific mRNAs during germ cell development.
CC -!- SUBUNIT: eIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least EIF4A, EIF4E and EIF4G. EIF4E is also
CC known to interact with other partners. The interaction with
CC EIF4ENIF1 mediates the import into the nucleus. Nonphosphorylated
CC EIF4B1, EIF4B2 and EIF4B3 compete with EIF4G to interact
CC with EIF4E; insulin stimulated MAP-kinase (MAPK1 and MAPK3)
CC phosphorylation of EIF4B1 causes dissociation of the complex
CC allowing EIF4G to bind and consequent initiation of translation.
CC Rapamycin can attenuate insulin stimulation, mediated by FKBPg.
CC -!- TISSUE SPECIFICITY: Very high levels in post-meiotic testicular
CC germ cells of rats of reproductive age.
CC -!- PTM: Phosphorylation increases the ability of the protein to bind
CC to mRNA caps and to form the EIF4F complex (By similarity).
CC -!- SIMILARITY: Belongs to the eukaryotic initiation factor 4E family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M61731; AAA37545.1; -;
DR EMBL; BC010759; AAH10759.1; -;
DR EMBL; X83399; CAA58316.1; -;
DR PIR; A34295; A34295.
DR PIR; I49644; I49644.
DR PDB; 1E34; 15-MAR-00.
DR PDB; 1EJH; 15-MAR-00.
DR PDB; 1L8B; 12-JUN-02.
DR MGD; MGI:95305; Eif4e.
DR InterPro; IPR001040; TIF_eIF_4E.
DR Pfam; PF01652; IF4E; 1.
DR ProDom; PD003697; TIF_eIF_4E; 1.
DR PROSITE; PS00813; IF4E; 1.
KW Protein biosynthesis; Translation regulation; Initiation factor;
KW RNA-binding; Phosphorylation; Multigene family; 3D-structure.
FT MOD_RES 209 209 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT CONFLICT 70 70 E -> L (IN REF. 1).
SQ SEQUENCE 217 AA; 25053 MW; FC61D0FE337BCD8F CRC64;

Query Match 83.9%; Score 26; DB 1; Length 217;
Best Local Similarity 50.0%; Pred. NO. 2.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
Db 43 WALWFF 48

RESULT 44
IF4E RABIT
ID IF4E RABIT STANDARD; PRT; 217 AA.
AC P29338;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Eukaryotic translation initiation factor 4E (eIF4E) (mRNA
DE cap-binding protein) (eIF-4F 25 kDa subunit).
GN EIF4E.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93117125; PubMed=1475206;
RA Rychlik W., Rhoads R.E.;

RT "Nucleotide sequence of rabbit eIF-4E cDNA.";
RL Nucleic Acids Res. 20:6415-6415(1992).
RN [2]
RP INTERACTION WITH E1F4A.
RX MEDLINE=83213553; PubMed=6853548;
RA Grifo J.A., Tahara S.M., Morgan M.A., Shatkin A.J., Merrick W.C.;
RT "New initiation factor activity required for globin mRNA
translation.";
RL J. Biol. Chem. 258:5804-5810(1983).
CC -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
mRNA cap during an early step in the initiation of protein
synthesis and facilitates ribosome binding by inducing the
unwinding of the mRNAs secondary structures.
CC -!- SUBUNIT: E1F4F is a multi-subunit complex, the composition of
which varies with external and internal environmental conditions.
CC It is composed of at least E1F4A, E1F4E and E1F4G. E1F4E is also
known to interact with other partners. The interaction with
E1F4ENIF1 mediates the import into the nucleus (By similarity).
CC -!- PTM: Phosphorylation increases the ability of the protein to bind
to mRNA caps and to form the E1F4 complex (By similarity).
CC -!- SIMILARITY: Belongs to the eukaryotic initiation factor 4E family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X61939; CAA43943.1; -.
DR PIR; S30248; B26411.
DR HSSP; P07260; 1AP8.
DR InterPro; IPR001040; TIF_eif_4E.
DR Pfam; PF01652; IF4E; 1.
DR ProDom; PD003697; TIF_eif_4E; 1.
DR PROSITE; PS00813; IF4E; 1.
KW Protein biosynthesis; Translation regulation; Initiation factor;
KW RNA-binding; Phosphorylation; Multigene family.
FT MOD_RES 209 209 PHOSPHORYLATION [BY PKC] (BY SIMILARITY).
SQ SEQUENCE 217 AA; 25049 MW; BSA6BE12F417159 CRC64;

Query Match 83.9%; Score 26; DB 1; Length 217;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6
Db 43 WALWFF 48

RESULT 45
YC07_METJA
ID YC07_METJA STANDARD; PRT; 226 AA.
AC Q58604;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical acetyltransferase MJ1207 (EC 2.3.1.-).
GN MJ1207.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Xlenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U67502; AAB98515.1; -.

RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Xlenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE ACETYLTRANSFERASE FAMILY. STRONG, TO
CC A.FULGIDUS AF0521.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U67562; AAB99211.1; -.
DR PIR; F64450; F64450.
DR TIGR; MJ1207; -.
DR InterPro; IPR000182; GCN5acetyl_trans.
DR Pfam; PF00583; Acetyltransf; 1.
KW Hypothetical protein; Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 226 AA; 26939 MW; 376E718D3509E2DA CRC64;
Query Match 83.9%; Score 26; DB 1; Length 226;
Best Local Similarity 50.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWKF 6
Db 109 WARWYF 114

RESULT 46
Y523_METJA
ID Y523_METJA STANDARD; PRT; 234 AA.
AC Q57943;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0523.
GN MJ0523.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Xlenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U67502; AAB98515.1; -.

Query Match 83.9% Score 26 DB 1: Length 234:

RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R.L., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium *Clostridium acetobutylicum*.";
 RL J. Bacteriol. 183:4823-4838(2001).
 CC -!- FUNCTION: Required for the insertion of integral membrane proteins
 CC into the membrane. Probably plays an essential role in the
 CC integration of proteins of the respiratory chain complexes.
 CC Involved in integration of membrane proteins that insert
 CC dependently and independently of the Sec translocase complex (By
 CC similarity).
 CC -!- SUBUNIT: Specifically interacts with transmembrane segments of
 CC nascent integral membrane proteins during membrane integration (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the OXA1/oxaA family. Subfamily 3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE007868; AAK81656.1; -;
 DR PIR; E97358; E97358.
 DR HAMAP; MF_01810; atypical; 1.
 DR InterPro; IPR001708; 60kDa_innermem.
 DR Pfam; PF02096; 60KD IMP; 1.
 KW Transmembrane; Complete proteome.
 FT TRANSMEM 31 53 POTENTIAL.
 FT TRANSMEM 98 120 POTENTIAL.
 FT TRANSMEM 141 163 POTENTIAL.
 FT TRANSMEM 183 205 POTENTIAL.
 SQ SEQUENCE 254 AA; 29017 MW; 4E2A0C6551D951A1 CRC64;
 Query Match 83.9%; Score 26; DB 1; Length 254;
 Best Local Similarity 50.0%; Pred. No. 3.3e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 Db 190 WMSWNP 195
 RESULT 50
 Y090 MYCTU STANDARD; PRT; 256 AA.
 AC Q10887;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein RV0090/MT0099/MB0093.
 GN RV0090 OR MT0099 OR MTCY251.08 OR MB0093.
 OS Mycobacterium tuberculosis, and
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 ON NCBI_taxid=1773, 1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.B. III, Tekala F.,
 RA Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkoeh;
 RX MEDLINE=2206494; PubMed=12218036;
 RA Fetschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL J. Bacteriol. 184:5479-5490(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.bovis; STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z74410; CAA98926.1; -;
 DR EMBL; AE006921; AAK44321.1; -;
 DR EMBL; BX248334; CAD92955.1; -;
 DR PIR; B70750; B70750.
 DR TIGR; MT0099; -;
 DR TubercuList; RV0090; -;
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 203 223 POTENTIAL.
 SQ SEQUENCE 256 AA; 27837 MW; 01033C21199DEC51 CRC64;
 Query Match 83.9%; Score 26; DB 1; Length 256;
 Best Local Similarity 50.0%; Pred. No. 3.3e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 Db 102 WAVNEF 107
 Search completed: June 10, 2004, 10:48:39
 Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:43:04 ; Search time 33 Seconds
(without alignments)
57.367 Million cell updates/sec

Title: US-09-912-414-9
Perfect score: 31
Sequence: 1 WXXWXF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 500 summaries

Database :

SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mnc:
8: sp_organellae:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	90.3	87	16	Q9CGR4 lactococcus
2	28	90.3	235	5	Q86AM7 dictyosteli
3	28	90.3	237	5	Q8T623 dictyosteli
4	28	90.3	328	16	Q9KUA4 vibrio chol
5	28	90.3	328	16	Q87LX8 vibrio para
6	28	90.3	398	16	Q9PQ84 ureaplasma
7	28	90.3	505	16	Q9JYF5 neisseria m
8	28	90.3	505	16	Q9JTF0 neisseria g
9	28	90.3	572	2	P72076 neisseria g
10	28	90.3	598	16	Q8F0P2 leptospira
11	28	90.3	707	5	Q95XL6 caenorhabdi
12	28	90.3	797	10	Q8SAY5 oryza sativ
13	27	87.1	28	6	Q62821 bubalus bub
14	27	87.1	92	5	Q8MQ59 caenorhabdi
15	27	87.1	109	16	Q8YW25 anabaena sp
16	27	87.1	135	6	Q95LT4 macaca fasc

17	87.1	137	13	O42234	coturnix co
18	87.1	139	16	Q92JV2	rhizobium m
19	87.1	141	11	Q8BS16	mus musculus
20	87.1	146	10	Q41182	nicotiana t
21	87.1	147	4	Q8N7Y7	homo sapien
22	87.1	151	16	Q55180	synecocyst
23	87.1	155	16	Q8X892	escherichia
24	87.1	176	9	Q37841	bacterioph
25	87.1	176	9	Q858V5	bacterioph
26	87.1	176	9	Q7Y4D5	bacterioph
27	87.1	186	16	Q92FG6	listeria in
28	87.1	186	16	Q8YAM4	listeria mo
29	87.1	214	5	O45341	caenorhabdi
30	87.1	216	2	Q8GM19	streptomyce
31	87.1	217	3	Q9CLK3	nectria hae
32	87.1	217	16	Q55732	synecocyst
33	87.1	221	16	Q9K1K8	neisseria m
34	87.1	221	16	Q9JWY7	neisseria m
35	87.1	227	12	Q9YS36	lactate deh
36	87.1	227	12	Q83019	lactate deh
37	87.1	227	12	Q9YS41	lactate deh
38	87.1	230	5	Q8WQM8	artemia san
39	87.1	239	10	Q9SL57	arabidopsis
40	87.1	244	5	Q95S28	drosophila
41	87.1	245	16	Q8X908	escherichia
42	87.1	245	16	Q8FEQ6	escherichia
43	87.1	245	16	Q83JZ6	shigella fl
44	87.1	247	16	Q829J1	streptomyce
45	87.1	255	17	Q8U324	pyrococcus
46	87.1	270	16	Q8F276	leptospira
47	87.1	274	17	Q8ZUD5	pyrobaculum
48	87.1	275	16	Q89X05	bradyrhizob
49	87.1	293	5	O17009	caenorhabdi
50	87.1	295	3	O94511	arabidopsis
51	87.1	297	16	Q9KDS0	schizosacch
52	87.1	301	10	Q8H3H8	oryza sativ
53	87.1	302	2	Q8GJ78	mycobacteri
54	87.1	307	11	Q8BGW7	mus musculus
55	87.1	318	16	Q8ZCU0	versinia pe
56	87.1	325	16	Q82XU0	streptomyce
57	87.1	333	5	Q9U346	caenorhabdi
58	87.1	348	11	Q9WTK1	cavia porce
59	87.1	357	17	Q97YW8	sulfolobus
60	87.1	378	5	Q95QB3	caenorhabdi
61	87.1	381	5	Q9TVV2	caenorhabdi
62	87.1	392	16	Q92MA2	rhizobium m
63	87.1	403	16	Q8DMH7	synecococc
64	87.1	406	11	Q91ZQ0	rattus norv
65	87.1	407	16	Q8P407	xanthomonas
66	87.1	414	5	Q966F1	caenorhabdi
67	87.1	424	16	Q7U3C9	synecococc
68	87.1	425	16	Q8PFI7	xanthomonas
69	87.1	431	5	Q19376	caenorhabdi
70	87.1	432	3	Q872N2	neurospora
71	87.1	434	13	O57404	pleurodeles
72	87.1	439	4	Q9BUG9	homo sapien
73	87.1	441	16	Q9HYE2	pseudomonas
74	87.1	449	16	Q8G4K2	bifidobacte
75	87.1	450	13	Q8AYB3	brachydanio
76	87.1	456	16	Q93JK1	streptomyce
77	87.1	469	16	Q98342	rhizobium l
78	87.1	495	16	Q8EGR9	shewanella
79	87.1	512	16	Q8ZMK8	salmonella
80	87.1	512	16	Q8ZAD9	ealmonella
81	87.1	520	16	Q8NRZ6	corynebacte
82	87.1	543	1	Q977V6	pyrococcus
83	87.1	550	12	O40912	kaposi's sa
84	87.1	579	16	Q81UM5	bacillus an
85	87.1	580	2	Q8RLZ8	acinetobact
86	87.1	596	16	Q81L51	bacillus an
87	87.1	628	16	Q9A4M9	caulobacter
88	87.1	647	16	Q8XRF0	ralstonia s
89	87.1	647	16	Q8XRF0	ralstonia s

236	26	83.9	255	2	Q93RD8	Q93rd8 thermochrom	309	26	83.9	323	16	Q7UDJ9	Q7udj9 shigella fl
237	26	83.9	255	16	Q82KL0	Q82k10 streptomyc	310	26	83.9	325	10	Q8LD77	Q8ld77 arabisopsis
238	26	83.9	256	10	Q9FIN2	Q9fin2 arabisopsis	311	26	83.9	325	10	Q8LD77	Q8ld77 arabisopsis
239	26	83.9	256	16	Q8NLA5	Q8nla5 corynebacte	312	26	83.9	331	10	Q84W2	Q84w2 arabisopsis
240	26	83.9	259	16	Q8ZPZ5	Q8zpz5 yersinia pe	313	26	83.9	331	10	Q84W2	Q84w2 arabisopsis
241	26	83.9	260	16	Q8GSD9	Q8gsd9 bifidobacte	314	26	83.9	331	16	Q8NRA3	Q8nra3 corynebacte
242	26	83.9	261	13	Q7SKW4	Q7skw4 brachydanio	315	26	83.9	333	16	Q8NRA3	Q8nra3 corynebacte
243	26	83.9	263	5	Q9VRM3	Q9vrn3 drosophila	316	26	83.9	334	5	Q9KWL7	Q9kwl7 corynebacte
244	26	83.9	266	2	Q88109	Q88109 mycobacteri	317	26	83.9	334	5	Q9KWL7	Q9kwl7 corynebacte
245	26	83.9	268	5	Q8MZE7	Q8mze7 drosophila	318	26	83.9	334	5	Q9KWL7	Q9kwl7 corynebacte
246	26	83.9	269	17	Q8ZY92	Q8zy92 pyrobaculum	319	26	83.9	334	16	Q8KG09	Q8kg09 chlorobium
247	26	83.9	270	16	Q8RM63	Q8rm63 deinococcus	320	26	83.9	334	16	Q8KG09	Q8kg09 chlorobium
248	26	83.9	272	16	Q54140	Q54140 streptomyc	321	26	83.9	334	16	Q8KG09	Q8kg09 chlorobium
249	26	83.9	274	16	Q8YQ88	Q8yq88 anabaena sp	322	26	83.9	337	11	Q8GDM1	Q8gdm1 heliobacill
250	26	83.9	276	16	Q8PQ1	Q8pq1 mycoplasma	323	26	83.9	337	11	Q8GDM1	Q8gdm1 heliobacill
251	26	83.9	276	17	Q8PUY8	Q8puy8 methanosarc	324	26	83.9	337	11	Q8GDM1	Q8gdm1 heliobacill
252	26	83.9	280	16	Q7U3L0	Q7u3l0 synchococc	325	26	83.9	337	11	Q8GDM1	Q8gdm1 heliobacill
253	26	83.9	281	5	Q85RQ4	Q85rq4 drosophila	326	26	83.9	338	16	Q83SG6	Q83sg6 shigella fl
254	26	83.9	281	16	Q8U554	Q8u554 agrobacteri	327	26	83.9	338	16	Q83SG6	Q83sg6 shigella fl
255	26	83.9	281	16	Q8U554	Q8u554 agrobacteri	328	26	83.9	338	16	Q83SG6	Q83sg6 shigella fl
256	26	83.9	281	16	Q7VR73	Q7vr73 candidatus	329	26	83.9	339	16	Q83SG6	Q83sg6 shigella fl
257	26	83.9	285	16	Q82ME6	Q82me6 streptomyc	330	26	83.9	339	16	Q83SG6	Q83sg6 shigella fl
258	26	83.9	285	16	Q7MN94	Q7mn94 bordetella	331	26	83.9	339	16	Q83SG6	Q83sg6 shigella fl
259	26	83.9	285	16	Q7MN94	Q7mn94 bordetella	332	26	83.9	339	16	Q83SG6	Q83sg6 shigella fl
260	26	83.9	286	16	Q9PDM4	Q9pdn4 xylella fas	333	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
261	26	83.9	286	16	Q8Y2G9	Q8y2g9 heliobacte	334	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
262	26	83.9	286	16	Q8PNN6	Q8pnn6 xanthomonas	335	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
263	26	83.9	286	16	Q8PC13	Q8pc13 xanthomonas	336	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
264	26	83.9	286	16	Q87DU1	Q87du1 xylella fas	337	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
265	26	83.9	287	16	Q8PHM5	Q8phm5 campylobact	338	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
266	26	83.9	287	16	Q8TRJ3	Q8trj3 methanosarc	339	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
267	26	83.9	289	17	Q8TRJ3	Q8trj3 methanosarc	340	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
268	26	83.9	290	10	Q94OM8	Q94om8 arabisopsis	341	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
269	26	83.9	290	16	Q8KAQ2	Q8kaq2 chlorobium	342	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
270	26	83.9	291	10	Q8LEY8	Q8ley8 arabisopsis	343	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
271	26	83.9	291	16	Q8PFT9	Q8pft9 xanthomonas	344	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
272	26	83.9	291	16	Q8P482	Q8p482 xanthomonas	345	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
273	26	83.9	291	16	Q8DH93	Q8dh93 synchococc	346	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
274	26	83.9	292	2	Q87196	Q87196 heliobacte	347	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
275	26	83.9	292	16	Q8ZMT3	Q8zmt3 heliobacte	348	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
276	26	83.9	295	10	Q8LA20	Q8la20 arabisopsis	349	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
277	26	83.9	297	16	Q9A4B6	Q9a4b6 caulobacter	350	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
278	26	83.9	300	16	Q91479	Q91479 pseudomonas	351	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
279	26	83.9	300	16	Q9HW34	Q9hw34 pseudomonas	352	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
280	26	83.9	301	16	Q929W0	Q929w0 listeria in	353	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
281	26	83.9	301	16	Q8Y5K3	Q8y5k3 listeria mo	354	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
282	26	83.9	305	16	Q83HJ2	Q83hj2 tropheryma	355	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
283	26	83.9	308	16	Q9L149	Q9l149 streptomyc	356	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
284	26	83.9	308	16	Q82PA3	Q82pa3 streptomyc	357	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
285	26	83.9	309	2	Q9KWB7	Q9kwb7 agrobacteri	358	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
286	26	83.9	311	16	Q7V3X3	Q7v3x3 prochloroco	359	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
287	26	83.9	312	16	Q8X764	Q8x764 escherichia	360	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
288	26	83.9	312	16	Q83IT2	Q83it2 shigella fl	361	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
289	26	83.9	314	16	Q9L120	Q9l120 streptomyc	362	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
290	26	83.9	314	16	Q83GP2	Q83gp2 tropheryma	363	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
291	26	83.9	314	16	Q82KA7	Q82ka7 streptomyc	364	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
292	26	83.9	315	10	Q9SW35	Q9sw35 arabisopsis	365	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
293	26	83.9	315	16	Q834R3	Q834r3 enterococc	366	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
294	26	83.9	316	5	Q62517	Q62517 caenorhabdi	367	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
295	26	83.9	316	16	Q88W05	Q88w05 lactobacill	368	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
296	26	83.9	317	2	Q9F9B8	Q9f9b8 paenibacill	369	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
297	26	83.9	317	10	Q9MAU4	Q9mau4 arabisopsis	370	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
298	26	83.9	317	16	Q9KE50	Q9ke50 bacillus ha	371	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
299	26	83.9	317	16	Q9X273	Q9x273 thermocoga	372	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
300	26	83.9	317	16	Q88OH0	Q88oh0 pseudomonas	373	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
301	26	83.9	318	16	Q87KC1	Q87kc1 vibrio para	374	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
302	26	83.9	319	2	Q53165	Q53165 renibacteri	375	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
303	26	83.9	320	17	Q9PVB5	Q9pnb5 aeropyrum p	376	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
304	26	83.9	321	16	Q06496	Q06496 clostridium	377	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
305	26	83.9	321	16	Q8FHS8	Q8fhs8 escherichia	378	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
306	26	83.9	321	16	Q83RL8	Q83rl8 shigella fl	379	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
307	26	83.9	322	10	Q9ZWF1	Q9zwf1 oryza sativ	380	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
308	26	83.9	323	5	Q81704	Q81704 trypanosoma	381	26	83.9	340	16	Q83SG6	Q83sg6 shigella fl
			323	16	Q8XGX9	Q8xgx9 salmonella							

382	26	83.9	378	8	Q9MNX0	Q9mnx0 pagurus lon	455	26	83.9	425	16	Q82VZ6	Q82vz6 nitrosomona
383	26	83.9	379	16	Q8G3E1	Q8g3el bruceella su	456	26	83.9	426	16	Q81NU1	Q81nu1 bacillus an
384	26	83.9	380	16	Q8YEE1	Q8yeel bruceella me	457	26	83.9	429	16	Q8Z508	Q8z508 salmonella
385	26	83.9	381	16	Q88JG9	Q88jg9 pseudomonas	458	26	83.9	430	2	Q9F9T3	Q9f9t3 edta-degrad
386	26	83.9	382	2	Q93TQ4	Q93tg4 agrobacteri	459	26	83.9	434	10	Q9LYL2	Q9lyl2 arabidopsis
387	26	83.9	383	8	Q36954	Q36954 chlamydomon	460	26	83.9	436	10	Q8LP16	Q8lp16 arabidopsis
388	26	83.9	384	8	Q8LWU0	Q8lww0 chlamydomon	461	26	83.9	437	16	Q92U00	Q92u00 rhizobium m
389	26	83.9	385	16	Q82LNO	Q82lno streptomyce	462	26	83.9	437	11	Q8VXZ7	Q8vxz7 arabidopsis
390	26	83.9	386	16	Q7UAH1	Q7uah1 shigella fl	463	26	83.9	437	11	Q35249	Q35249 rattus norv
391	26	83.9	387	16	Q8ZQD2	Q8zqd2 salmonella	464	26	83.9	437	16	Q7V3T8	Q7v3t8 prochloroco
392	26	83.9	388	16	Q8XEB7	Q8xeb7 escherichia	465	26	83.9	438	8	Q8LYW1	Q8lyw1 polytomella
393	26	83.9	389	16	Q8RKT2	Q8rkt2 streptomyce	466	26	83.9	439	16	Q9K5R3	Q9k5r3 bacillus ha
394	26	83.9	390	16	Q8RKT2	Q8rkt2 streptomyce	467	26	83.9	442	16	Q05224	Q05224 bacillus su
395	26	83.9	391	16	Q8FJC1	Q8fjc1 escherichia	468	26	83.9	446	16	Q05224	Q05224 bacillus su
396	26	83.9	392	16	Q82810	Q82810 salmonella	469	26	83.9	448	16	Q9K5W2	Q9k5w2 bradyrhizob
397	26	83.9	393	16	Q83RZ5	Q83rz5 shigella fl	470	26	83.9	448	16	Q9K5W2	Q9k5w2 bacillus ha
398	26	83.9	394	16	Q8CT10	Q8ct10 staphylococ	471	26	83.9	451	10	Q9C9U6	Q9c9u6 arabidopsis
399	26	83.9	395	16	Q8TKD9	Q8tkd9 methanosarc	472	26	83.9	452	17	Q26414	Q26414 methanobact
400	26	83.9	396	16	Q8UHG1	Q8uhg1 rhizobium l	473	26	83.9	454	10	Q9ZRW0	Q9zrw0 cicer ariet
401	26	83.9	397	4	Q8UHG1	Q8uhg1 rhizobium l	474	26	83.9	457	16	Q8F602	Q8f602 leptospira
402	26	83.9	398	2	Q8N7C3	Q8n7c3 homo sapien	475	26	83.9	458	16	Q45613	Q45613 bacillus su
403	26	83.9	399	10	Q7XMS7	Q7xms7 oryza sativ	476	26	83.9	460	8	Q9T910	Q9t910 megalobrama
404	26	83.9	400	17	Q8TZ67	Q8tz67 methanopyru	477	26	83.9	460	10	Q9STI4	Q9sti4 arabidopsis
405	26	83.9	401	16	Q89DN7	Q89dn7 bradyrhizob	478	26	83.9	461	8	Q33748	Q33748 arabacia lix
406	26	83.9	402	16	Q89DN7	Q89dn7 bradyrhizob	479	26	83.9	461	5	Q18910	Q18910 caenorhabdi
407	26	83.9	403	16	Q8YF94	Q8yfu5 bruceella me	480	26	83.9	463	16	Q7VMB5	Q7vmb5 haemophilus
408	26	83.9	404	16	Q8YF94	Q8yfu5 bruceella me	481	26	83.9	465	16	Q8EXJ8	Q8exj8 leptospira
409	26	83.9	405	16	Q8FZ16	Q8fz16 bruceella su	482	26	83.9	468	16	Q8ZNC7	Q8znc7 salmonella
410	26	83.9	406	16	Q8UJ08	Q8uj08 agrobacteri	483	26	83.9	468	16	Q8YT65	Q8yt65 anabaena sp
411	26	83.9	407	10	Q9AW80	Q9aw80 guillardia	484	26	83.9	469	16	Q7UJ09	Q7ujm9 rhodopirell
412	26	83.9	408	16	Q8A119	Q8a119 bacteroides	485	26	83.9	469	17	Q8PYA6	Q8pya6 methanosarc
413	26	83.9	409	16	Q8A119	Q8a119 bacteroides	486	26	83.9	471	5	Q23488	Q23488 caenorhabdi
414	26	83.9	410	2	Q59719	Q59719 pseudomonas	487	26	83.9	471	10	Q8H3A6	Q8h3a6 oryza sativ
415	26	83.9	411	16	Q9A3I8	Q9a3i8 caulobacter	488	26	83.9	473	11	Q9Z1X2	Q9z1x2 mus musculu
416	26	83.9	412	16	Q929D7	Q929d7 listeria in	489	26	83.9	473	11	Q8BHL2	Q8bhl2 mus musculu
417	26	83.9	413	16	Q8Y538	Q8y538 listeria in	490	26	83.9	474	11	Q08888	Q08888 cricetus
418	26	83.9	414	16	Q8Y538	Q8y538 listeria in	491	26	83.9	475	16	Q8DY99	Q8dy99 streptococ
419	26	83.9	415	16	Q7WE49	Q7we49 bordetella	492	26	83.9	475	16	Q9ABD8	Q9abd8 caulobacter
420	26	83.9	416	16	Q7VS06	Q7vso6 bordetella	493	26	83.9	477	16	Q8DPS3	Q8dps3 streptococ
421	26	83.9	417	16	Q922A1	Q922a1 mus musculu	494	26	83.9	479	16	Q97SA7	Q97sa7 streptococ
422	26	83.9	418	16	Q8NS00	Q8ns00 corynebacte	495	26	83.9	479	16	Q8E3W4	Q8e3w4 streptococ
423	26	83.9	419	16	Q83RA4	Q83ra4 shigella fl	496	26	83.9	479	16	Q8DSR2	Q8dsr2 streptococ
424	26	83.9	420	16	Q9A6G3	Q9a6g3 caulobacter	497	26	83.9	479	16	Q8DY99	Q8dy99 streptococ
425	26	83.9	421	16	Q8ZPR2	Q8zpr2 salmonella	498	26	83.9	481	16	Q7V8Y2	Q7v8y2 prochloroco
426	26	83.9	422	16	Q8YFS6	Q8yfs6 bruceella me	499	26	83.9	484	16	Q8F4I3	Q8f4i3 leptospira
427	26	83.9	423	5	Q21147	Q21147 caenorhabdi	500	26	83.9	484	16	Q88VK9	Q88vk9 lactobacill
428	26	83.9	424	16	Q8XSY5	Q8xsy5 escherichia							
429	26	83.9	425	16	Q8FH44	Q8fh44 escherichia							
430	26	83.9	426	16	Q8UGM3	Q8ugm3 agrobacteri							
431	26	83.9	427	16	Q8H0P4	Q8h0p4 homo sapien							
432	26	83.9	428	4	Q86GC9	Q86gc9 homo sapien							
433	26	83.9	429	11	Q99K00	Q99ku0 mus musculu							
434	26	83.9	430	11	Q8BHD3	Q8bhd3 mus musculu							
435	26	83.9	431	16	Q89UJ0	Q89uj0 bradyrhizob							
436	26	83.9	432	16	Q89UJ0	Q89uj0 bradyrhizob							
437	26	83.9	433	16	Q89UJ0	Q89uj0 bradyrhizob							
438	26	83.9	434	16	Q89UJ0	Q89uj0 bradyrhizob							
439	26	83.9	435	16	Q89UJ0	Q89uj0 bradyrhizob							
440	26	83.9	436	16	Q89UJ0	Q89uj0 bradyrhizob							
441	26	83.9	437	16	Q89UJ0	Q89uj0 bradyrhizob							
442	26	83.9	438	16	Q89UJ0	Q89uj0 bradyrhizob							
443	26	83.9	439	16	Q89UJ0	Q89uj0 bradyrhizob							
444	26	83.9	440	16	Q89UJ0	Q89uj0 bradyrhizob							
445	26	83.9	441	16	Q89UJ0	Q89uj0 bradyrhizob							
446	26	83.9	442	16	Q89UJ0	Q89uj0 bradyrhizob							
447	26	83.9	443	16	Q89UJ0	Q89uj0 bradyrhizob							
448	26	83.9	444	16	Q89UJ0	Q89uj0 bradyrhizob							
449	26	83.9	445	16	Q89UJ0	Q89uj0 bradyrhizob							
450	26	83.9	446	16	Q89UJ0	Q89uj0 bradyrhizob							
451	26	83.9	447	16	Q89UJ0	Q89uj0 bradyrhizob							
452	26	83.9	448	16	Q89UJ0	Q89uj0 bradyrhizob							
453	26	83.9	449	16	Q89UJ0	Q89uj0 bradyrhizob							
454	26	83.9	450	16	Q89UJ0	Q89uj0 bradyrhizob							

ALIGNMENTS

RESULT 1	ID	Q9CGR4	PRELIMINARY;	PRT;	87 AA.
Q9CGR4	AC	Q9CGR4;			
DT	01-JUN-2001	(TrEMBLrel. 17, Created)			
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)			
DE	Hypothetical protein p1225.				
GN	P1225 OR L11032.				
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.				
OX	NCBI_taxID=1360;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=IL1403;				
RC	MEDLINE=21235186; PubMed=11337471;				
RA	Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,				
RA	Weissenbach J., Ehrlich S.D., Sorokin A.;				
RT	"The complete genome sequence of the lactic acid bacterium Lactococcus				
RT	lactis ssp. lactis Il1403."				
RL	Genome Res. 11:731-753(2001).				
DR	EMBL; AE006336; AAK05130.1; -.				

```

DR PIR; H86753; H86753.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 87 AA; 10178 MW; E7A0A0AA3D08E6D3 CRC64;

Query Match          90.3%; Score 28; DB 16; Length 87;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
Db 17 WSAAWF 22

RESULT 2
Q86AM7 PRELIMINARY; PRT; 235 AA.
AC Q86AM7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Similar to Dictyostelium discoideum (Slime mold). prenyl cysteine
DE carboxyl methyltransferase.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA "Segal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116984; AA051329.1; -.
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR007269; ICMT.
DR Pfam; PF04140; ICMT; 1.
KW Methyltransferase; Transferase.
SQ SEQUENCE 235 AA; 27177 MW; CDCAA6C9AF85014 CRC64;

Query Match          90.3%; Score 28; DB 5; Length 235;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
Db 195 WASWSF 200

RESULT 3
Q8T623 PRELIMINARY; PRT; 237 AA.
AC Q8T623;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Prenyl cysteine carboxyl methyltransferase.
OS Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Y., Stock J.B., Cox E.C.;
RT "Prenyl Cysteine Carboxyl Methyltransferase in Dictyostelium.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF487784; AAL99548.1; -.

```

```

DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR007269; ICMT.
DR Pfam; PF04140; ICMT; 1.
KW Transferase; Methyltransferase.
SQ SEQUENCE 237 AA; 27379 MW; 3F6961685B01C5DE CRC64;

Query Match          90.3%; Score 28; DB 5; Length 237;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
Db 197 WASWSF 202

RESULT 4
Q9KUA4 PRELIMINARY; PRT; 328 AA.
AC Q9KUA4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Peptide ABC transporter, permease protein.
GN VC0619.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Bisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483 (2000).
DR PIR; G82301; G82301.
DR TIGR; VC0619; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp; 1.
KW Complete proteome.
SQ SEQUENCE 328 AA; 36261 MW; 210127776BA69196 CRC64;

Query Match          90.3%; Score 28; DB 16; Length 328;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
Db 180 WTSWAF 185

RESULT 5
Q87LX8 PRELIMINARY; PRT; 328 AA.
AC Q87LX8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Peptide ABC transporter, permease protein.
GN VP2480.
OS Vibrio parahaemolyticus.

```

Tue Jun 15 10:57:37 2004

us-09-912-414-9-rspt

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=670;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism
 distinct from that of *V. cholerae*.";
 RL Lancet 361:743-749 (2003).
 DR EMBL; AP005081; BAC60743.1; --
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 KW Complete proteome.
 SQ SEQUENCE 328 AA; 36196 MW; FB4704C070B55DC3 CRC64;
 Query Match 90.3%; Score 28; DB 16; Length 328;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 Db 180 WSSWAF 185
 RESULT 6
 Q9PQ84 PRELIMINARY; PRT; 398 AA.
 AC Q9PQ84
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein U0405.1.
 GN U0405.1.
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=134821;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serovar 3;
 RX MEDLINE=20500219; PubMed=11048724;
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
 RA Cassell G.H.;
 RT "The complete sequence of the mucosal pathogen *Ureaplasma*
urealyticum.";
 RL Nature 407:757-762 (2000).
 DR EMBL; AE002138; AAF30816.1; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 398 AA; 47349 MW; 90C7882DD9B42FFC CRC64;
 Query Match 90.3%; Score 28; DB 16; Length 398;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 Db 135 WTSWSF 140
 RESULT 7
 Q9JYF5 PRELIMINARY; PRT; 505 AA.
 AC Q9JYF5
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sensor histidine kinase.

GN NMB1606.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Neftci W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Cittiore H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
 MC58.";
 RL Science 287:1809-1815 (2000).
 CC -!- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
 DR EMBL; AE002510; AAF41958.1; --
 DR PIR; H81064; H81064.
 DR TIGR; NMB1606; --
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO; GO:0007600; P:sensory perception; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR005467; His_kinase.
 DR InterPro; IPR003661; His_kinase_N.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF00512; HSKA; 1.
 DR SMART; SM00388; HSKA; 1.
 DR PROSITE; PS01019; HIS_KIN; 1.
 KW Kinase; Phosphorylation; Sensory transduction; Transferase;
 KW Complete proteome.
 SQ SEQUENCE 505 AA; 56085 MW; 97F13DCD75A748B6 CRC64;
 Query Match 90.3%; Score 28; DB 16; Length 505;
 Best Local Similarity 50.0%; Pred. No. 1.9e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 Db 51 WTAWAF 56
 RESULT 8
 Q9JTF0 PRELIMINARY; PRT; 505 AA.
 AC Q9JTF0
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative two component sensor kinase.
 GN NMB1803.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,

```

RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506 (2000).
CC -!- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL; AL162757; CAB85029.1; -.
DR F1R; A81806; A81806.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; AtPbind ATPase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF02518; HATPase_c7_1.
DR Pfam; PF00512; Hiska; 1.
DR SMART; SM00388; Hiska; 1.
DR PROSITE; PS01019; HIS_KIN; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase;
KW Complete proteome.
SQ SEQUENCE 505 AA; 56509 MW; 6CAAF8758E7AB186 CRC64;

Query Match 90.3%; Score 28; DB 16; Length 505;
Best Local Similarity 50.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 51 WTAWAF 56

RESULT 9
P72076 PRELIMINARY; PRT; 572 AA.
AC P72076;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RSP.
GN RSP.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
[1]
SEQUENCE FROM N.A.
RA STRAIN=MS11-A;
RC MEDLINE=98440364; PubMed=9767087;
RX Carrick C.S., Fyfe J.A.M., Davies J.K.;
RT "Neisseria gonorrhoeae contains multiple copies of a gene that may
RT encode a site-specific recombinase and is associated with DNA
RT rearrangements."
RL Gene 220:21-29 (1998).
CC -!- SIMILARITY: CONTAINS A SIGMA-54 FACTOR INTERACTION ATP-BINDING
CC DOMAIN.
DR EMBL; U65994; AAC82507.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR008931; FIS-like.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR002197; HTH_Fis_.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF02954; HTH_8; 1.
DR Pfam; PF00158; Sigma54_activat; 1.
DR PRINTS; PR01590; HTHFIS.
DR SMART; SM00388; Hiska; 1.

DR TIGRfams; TIGR01199; HTH_fis; 1.
DR PROSITE; PS00676; SIGMAS4_INTERACT_2; 1.
DR PROSITE; PS00445; SIGMAS4_INTERACT_4; 1.
KW ATP-binding; DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 572 AA; 64646 MW; A013C5BE0E516615 CRC64;

Query Match 90.3%; Score 28; DB 2; Length 572;
Best Local Similarity 50.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 51 WTAWAF 56

RESULT 10
Q8F0P2 PRELIMINARY; PRT; 598 AA.
ID Q8F0P2;
AC Q8F0P2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Regulatory protein, putative.
GN LA3450.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011502; AAN50648.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR001932; PP2C-like.
DR SMART; SM00331; PP2C_SIG; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 598 AA; 68794 MW; 1BF823A219B68A0F CRC64;

Query Match 90.3%; Score 28; DB 16; Length 598;
Best Local Similarity 50.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 60 WTAWAF 65

RESULT 11
Q95XL6 PRELIMINARY; PRT; 707 AA.
ID Q95XL6;
AC Q95XL6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN V73E7A.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018 (1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du H., Maupin R.;

```



```

RT "The sequence of C. elegans cosmid Y73E7A.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025727; AK66029.1; -.
DR WormPep; Y73E7A.8; CE27538.
DR InterPro; IPR000884; TSP1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00092; TSP1; 1.
KW Hypothetical protein.
SQ SEQUENCE 707 AA; 79940 MW; 26B924988153DBA8 CRC64;

Query Match 90.3%; Score 28; DB 5; Length 707;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 65 WSAWSF 70

RESULT 12
OBSAY5 PRELIMINARY; PRT; 797 AA.
AC Q8AY5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNBAA0010E04.17.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Heilao J., Ziemann V., Blunt S., Pai G.,
RA Vanaken S.E., Utkerback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.N.;
RT "Oryza sativa chromosome 3 BAC OSJNBAA0010E04 genomic sequence.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC096687; AAL79752.1; -.
DR Gramene; OBSAY5; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001480; B:lectin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR Pfam; PF01453; Agglutinin; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00108; B:lectin; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 797 AA; 87691 MW; D1812674CC1DC0F7 CRC64;

Query Match 90.3%; Score 28; DB 10; Length 797;
Best Local Similarity 50.0%; Pred. No. 2.7e+03;

RT "The sequence of C. elegans cosmid Y73E7A.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025727; AK66029.1; -.
DR WormPep; Y73E7A.8; CE27538.
DR InterPro; IPR000884; TSP1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00092; TSP1; 1.
KW Hypothetical protein.
SQ SEQUENCE 707 AA; 79940 MW; 26B924988153DBA8 CRC64;

Query Match 87.1%; Score 27; DB 6; Length 28;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
DB 10 WSVWTF 15

RESULT 14
QBMQ59 PRELIMINARY; PRT; 92 AA.
AC Q8MQ59;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein F35H12.6.
GN F35H12.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nelson J., Gattung S.;
RT "The sequence of C. elegans cosmid F35H12.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41540; AAM54184.1; -.

```

```

DR WormPep; F35H12.6; CE30973.
KW Hypothetical protein.
SQ SEQUENCE 92 AA; 10288 MW; 652728BC70B3D352 CRC64;
  Query Match      87.1%; Score 27; DB 5; Length 92;
  Best Local Similarity 50.0%; Pred. No. 6.8e+02;
  Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
  |||
Db 34 WATWLF 39

RESULT 15
ID Q8YW25 PRELIMINARY; PRT; 109 AA.
AC Q8YW25;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Alr1791.
GN ALR1791.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103630;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsumoto A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003597; BAB73430.1; -.
DR PIR; A12029; A12029.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 109 AA; 12167 MW; 6DED65C096A5C7B8 CRC64;

  Query Match      87.1%; Score 27; DB 16; Length 109;
  Best Local Similarity 50.0%; Pred. No. 7.8e+02;
  Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
  |||
Db 24 WATWLF 29

RESULT 16
ID Q95LT4 PRELIMINARY; PRT; 135 AA.
AC Q95LT4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071108; BAB64502.1; -.
KW Hypothetical protein.

SQ SEQUENCE 135 AA; 15499 MW; 23182747E172A019 CRC64;
  Query Match      87.1%; Score 27; DB 6; Length 135;
  Best Local Similarity 50.0%; Pred. No. 9.3e+02;
  Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
  |||
Db 60 WSWWLF 65

RESULT 17
ID O42234 PRELIMINARY; PRT; 137 AA.
AC O42234;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sonic hedgehog (Fragment).
GN SHH.
OS Coturnix coturnix (Common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98167903; PubMed=9435297;
RA Borycki A.G., Mendham L., Emerson C.P. Jr.;
RT "Control of somite patterning by Sonic hedgehog and its downstream
RT signal response genes.";
RL Development 125:777-790(1998).
DR EMBL; AF022882; AAB80949.1; -.
DR MEROPS; C46.002; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0007154; P:cell communication; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003586; Hedgehog_hintC.
DR InterPro; IPR001657; Peptidase_C46.
DR InterPro; IPR001767; Pept_C46_hint.
DR Pfam; PF01079; Hint; 1.
DR PRINTS; PR00632; SONICHOG.
DR SMART; SM00305; HintC; 1.
FT NON_TER 1
SQ SEQUENCE 137 AA; 14571 MW; 7FD29DF815AF1532 CRC64;

  Query Match      87.1%; Score 27; DB 13; Length 137;
  Best Local Similarity 50.0%; Pred. No. 9.4e+02;
  Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
  |||
Db 73 WAWWAF 78

RESULT 18
ID Q92JV2 PRELIMINARY; PRT; 139 AA.
AC Q92JV2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical transmembrane protein SMC02515.
GN R03020 OR SMC02515.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;

```

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Goudie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rampeger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT *Sinorhizobium meliloti* strain 1021.",
 RL Proc. Natl. Acad. Sci. U.S.A. 98:987-9882 (2001).
 DR EMBL; AL591792; CAC47599.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 139 AA; 15321 MW; 4B2627A5FE0970ED CRC64;

Query Match 87.1%; Score 27; DB 16; Length 139;
 Best Local Similarity 50.0%; Pred. No. 9.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXF 6
 Db 51 WMAWSF 56

RESULT 19
 Q8BS16 PRELIMINARY; PRT; 141 AA.
 ID Q8BS16
 AC Q8BS16
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Tripeptidyl peptidase II.
 OS *Mus musculus* (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.",
 RL Nature 420:563-573 (2002).
 DR EMBL; AK040859; BAC30721.1; -.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR00209; peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 141 AA; 14916 MW; 3A2CC953D78BE709 CRC64;

Query Match 87.1%; Score 27; DB 11; Length 141;
 Best Local Similarity 50.0%; Pred. No. 9.6e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXF 6
 Db 100 WSAWLF 105

RESULT 20
 Q41182 PRELIMINARY; PRT; 146 AA.
 ID Q41182
 AC Q41182
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Orf protein.
 GN ORF.
 OS *Nicotiana tabacum* (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=92404740; PubMed=2152343;
 RA Neale A.D., Walther J.A., Lund M., Bonnett H.T., Kelly A.,
 RA Meeks-Wagner D.R., Peacock W.J., Dennis E.S.;
 RT "Chitinase, beta-1,3-glucanase, osmotin, and extensin are expressed in
 RT tobacco explants during flower formation.",
 RL Plant Cell 2:673-684 (1990).
 DR EMBL; S44872; AAB33378.1; -.
 DR PIR; JQ0995; JQ0995.
 DR InterPro; IPR000916; Bet v I.
 DR Pfam; PF00407; Bet v I_1.
 SQ SEQUENCE 146 AA; 16469 MW; 7FC3EB16DC713D5C CRC64;

Query Match 87.1%; Score 27; DB 10; Length 146;
 Best Local Similarity 50.0%; Pred. No. 9.9e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXF 6
 Db 108 WMTWTF 113

RESULT 21
 Q8N7Y7 PRELIMINARY; PRT; 147 AA.
 ID Q8N7Y7
 AC Q8N7Y7
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ40209.
 OS *Homo sapiens* (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
 RA Osuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saico K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
 RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RT "NEDO human cDNA sequencing project.",
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK097528; BAC05089.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 147 AA; 16152 MW; 6D95546F07A4CED2 CRC64;

Query Match 87.1%; Score 27; DB 4; Length 147;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXF 6
 Db 50 WSTWRF 55

RESULT 22
 Q55180 PRELIMINARY; PRT; 151 AA.
 ID Q55180
 AC Q55180
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein slr0489.
 GN SLR0489.
 OS *Synechocystis* sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCC 6803;

RA Tabata S.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugiura M., Tabata S.;
 RA "Sequence analysis of the genome of the unicellular cyanobacterium
 RT *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 64% to 92% of the genome.";
 RL DNA Res. 2:153-166(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugitani M., Sasamoto S., Kimura T.,
 RA Hoshouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RA "Sequence analysis of the genome of the unicellular cyanobacterium
 RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D64001; BAA10320.1; -;
 DR PIR; S74402; S74402.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 151 AA; 17517 MW; EC46A9F3AE2A9122 CRC64;
 QY Query Match 87.1%; Score 27; DB 16; Length 151;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 WXXWKF 6
 DB 44 WSWWAF 49

RESULT 23
 QX8892 PRELIMINARY; PRT; 155 AA.
 ID QX8892
 AC QX8892;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative leader peptidase (Leader peptidase HOPD).
 GN HOPD OR Z4693 OR ECS4186.
 OS *Escherichia coli* O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RA "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).

DR EMBL; AB005557; AAC58443.1; -;
 DR EMBL; AP002564; BAB37611.1; -;
 DR PIR; D91152; D91152.
 DR PIR; G85997; G85997.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000045; Peptidase_A24A.
 DR Pfam; PF01478; Peptidase_A24; 1.
 DR PRINTS; PR00864; PREPILNPTASE.
 KW Complete proteome.
 SQ SEQUENCE 155 AA; 17099 MW; F65112397B336AD3 CRC64;
 QY Query Match 87.1%; Score 27; DB 16; Length 155;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 WXXWKF 6
 DB 96 WSWWAF 101

RESULT 24
 Q37841 PRELIMINARY; PRT; 176 AA.
 ID Q37841
 AC Q37841;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Orf38.
 OS Bacteriophage 186.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC P2-like viruses.
 OX NCBI_TaxID=29252;
 RN [1]
 RP SEQUENCE OF 174-176 FROM N.A.
 RX MEDLINE=95407087; PubMed=7676622;
 RA Xue Q., Egan J.B.;
 RT "DNA sequence of tail fiber genes of coliphage 186 and evidence for a
 RT common ancestor shared by dsDNA phage fiber genes.";
 RL Virology 212:128-133(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Xue Q.;
 RT "Studies on the tail region of the temperate coliphage 186 genome.";
 RL Thesis (1993), University of Adelaide.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98371265; PubMed=9705261;
 RA Portelli R., Dodd I.B., Xue Q., Egan J.B.;
 RT "The late-expressed region of the temperate coliphage 186 genome.";
 RL Virology 248:117-130(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Dodd I.B., Egan J.B.;
 RT "Bacteriophage 186 complete genome.";
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Xue Q., Egan J.B.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U32222; AAC34163.1; -;
 DR InterPro; IPR006521; Tail_P2_I.
 DR TIGRFAMs; TIGR01634; tail_P2_I; 1.
 SQ SEQUENCE 176 AA; 19643 MW; E53F5764AB650182 CRC64;
 QY Query Match 87.1%; Score 27; DB 9; Length 176;
 Best Local Similarity 50.0%; Pred. No. 1.2e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 WXXWKF 6
 DB 48 WSWWAF 53

```

OC Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596163; CAC95373.1; -.
DR PIR; AE1450; AE1450.
DR LIPID; LIN00140; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 186 AA; 21548 MW; 65D1B90CE40FFBF1 CRC64;

Query Match 87.1%; Score 27; DB 16; Length 186;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
Db 93 WSWLWF 98

RESULT 28
Q8YAM4 PRELIMINARY; PRT; 186 AA.
AC Q8YAM4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein lmo0094.
GN LMO0094.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591973; CAC98309.1; -.
DR PIR; AG1086; AG1086.
DR LIPID; LMO0094; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 186 AA; 21677 MW; 258A0D394B300ECB CRC64;

Query Match 87.1%; Score 27; DB 16; Length 186;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
Db 93 WSWLWF 98

OC Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596163; CAC95373.1; -.
DR PIR; AE1450; AE1450.
DR LIPID; LIN00140; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 186 AA; 21548 MW; 65D1B90CE40FFBF1 CRC64;

Query Match 87.1%; Score 27; DB 9; Length 176;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
Db 48 WLAWAF 53

RESULT 26
Q7Y4D5 PRELIMINARY; PRT; 176 AA.
AC Q7Y4D5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE GpI.
OS Bacteriophage Wphi.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=103216;
RN [1]
RP SEQUENCE FROM N.A.
RC Eposito D., Schmidt B.J., Bloom F.R., Christie G.E.;
RT "The Complete Nucleotide Sequence and Genome Analysis of the
RT Escherichia coli Bacteriophage Wphi.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV135739; AAN28236.1; -.
SQ SEQUENCE 176 AA; 19724 MW; 41A6F9DC749CFD8F CRC64;

Query Match 87.1%; Score 27; DB 9; Length 176;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
Db 48 WLAWAF 53

RESULT 27
Q92FG6 PRELIMINARY; PRT; 186 AA.
AC Q92FG6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein lin0140.
GN LIN0140.
OS Listeria innocua.

```

QY 1 WXXWXP 6
| | | |
Db 93 WSWLWF 98

RESULT 29

ID O45341 PRELIMINARY; PRT; 214 AA.
AC O45341;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F10A3.7 protein.
GN F10A3.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Rhabditinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z92829; CAB07346.1; -.
DR PIR; T20683; T20683.
DR WormPep; F10A3.7; CE15773.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR InterPro; IPR003002; 7TM_chemreceptl.
DR InterPro; IPR00168; Nm/TM_chemreceptl.
DR Pfam; PF01461; 7tm_4; 1.
SQ SEQUENCE 214 AA; 25088 MW; 7B73BC92B7E1D0F5 CRC64;

Query Match 87.1%; Score 27; DB 5; Length 214;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | | |
Db 6 WSWLWF 11

RESULT 30

ID Q8GM19 PRELIMINARY; PRT; 216 AA.
AC Q8GM19;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hydroxylase-like protein.
OS Streptomyces globisporus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1908;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C-1027;
RX MEDLINE=22171413; PubMed=12183628;
RA Liu W., Christenson S.D., Standage S., Shen B.;
RT "Biosynthesis of the enediyne antitumor antibiotic C-1027.";
RL Science 297:1170-1173(2002).
DR EMBL; AY048670; AAL06648.1; -.
SQ SEQUENCE 216 AA; 23976 MW; C0702A1A0375620B CRC64;

Query Match 87.1%; Score 27; DB 2; Length 216;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | | |
Db 97 WASWKF 102

RESULT 31

ID Q9CIK3 PRELIMINARY; PRT; 217 AA.
AC Q9CIK3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNA transposase.
OS Nectria haematococca.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.
OX NCBI_TaxID=140110;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=77-13-7;
RL Han Y., Liu X., Benny U., Kistler C.H., VanEtten H.D.;
RT "Genes determining pathogenicity to pea are clustered on a
RT supernumerary chromosome in the fungal plant pathogen, Nectria
RT haematococca.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF315315; AAK16923.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR InterPro; IPR002492; Transposase_5.
DR Pfam; PF01498; transposase_5; 1.
SQ SEQUENCE 217 AA; 25432 MW; 20C44094C212F8C8 CRC64;

Query Match 87.1%; Score 27; DB 3; Length 217;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
| | | |
Db 143 WSWTWF 148

RESULT 32

ID Q55732 PRELIMINARY; PRT; 217 AA.
AC Q55732;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein sll0397.
GN SLL0397.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PCC 6803;
RA Tabata S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneo T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugitara M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;

```

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asanizu E., Nakamura Y.,
RA Miyajima N., Hirotsawa M., Sugitara M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Mada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D64002; BAA10374.1; -.
DR PIR; S76528; S76528.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 217 AA; 25486 MW; 3A3F71A49DA035FE CRC64;

Query Match 87.1%; Score 27; DB 16; Length 217;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
DB 128 WSAWIF 133

RESULT 33
Q9K1K8 PRELIMINARY; PRT; 221 AA.
ID Q9K1K8
AC Q9K1K8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein NMB0108.
GN NMB0108.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; Pubmed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citron H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002369; AAP40567.1; -.
DR PIR; C81238; C81238.
DR TTGR; NMB0108; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 221 AA; 24705 MW; 52467C9FFE7FF8AC CRC64;

Query Match 87.1%; Score 27; DB 16; Length 221;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
DB 54 WLAWAF 59

RESULT 34
Q9JWY7 PRELIMINARY; PRT; 221 AA.
ID Q9JWY7
AC Q9JWY7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

```

DE Putative integral membrane protein.
GN NMA0166.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; Pubmed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162752; CAB83480.1; -.
DR PIR; C82010; C82010.
KW Complete proteome.
SQ SEQUENCE 221 AA; 24723 MW; 54867E5FFA7FF8AC CRC64;

Query Match 87.1%; Score 27; DB 16; Length 221;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
DB 54 WLAWAF 59

RESULT 35
Q9YS36 PRELIMINARY; PRT; 227 AA.
ID Q9YS36
AC Q9YS36;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Minor structural glycoprotein.
OS Lactate dehydrogenase-elevating virus (LDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=11048;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LDV-vx;
RX MEDLINE=99263186; Pubmed=10329569;
RA Li K., Chen Z., Plagemann P.;
RT "High-frequency homologous genetic recombination of an arterivirus,
RT lactate dehydrogenase-elevating virus, in mice and evolution of
RT neuropathogenic variants.";
RL Virology 258:73-83(1999).
DR EMBL; AF092283; AAC84057.1; -.
DR InterPro; IPR003434; PRSV.Env.
DR Pfam; PF02340; PRSV.Env; I.
SQ SEQUENCE 227 AA; 25914 MW; 96C579035AD98794 CRC64;

Query Match 87.1%; Score 27; DB 12; Length 227;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXF 6
DB 14 WFSWAF 19

RESULT 36
Q83019 PRELIMINARY; PRT; 227 AA.
ID Q83019
AC Q83019;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```

DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)	
DE	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)	
DE	Minor structural glycoprotein.		
OS	Lactate dehydrogenase-elevating virus (LDV).		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;		
OC	Arteriviridae; Arterivirus.		
OC	NCBI_TaxID=11048;		
OC	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=Plagemann;		
RC	MEDLINE=95297165; PubMed=7778295;		
RA	Palmer G.A., Kuo L., Chen Z., Faaberg K.S., Plagemann P.G.;		
RX	"Sequence of the genome of lactate dehydrogenase-elevating virus:		
RT	heterogeneity between strains P and C.,"		
RT	Virology 209:637-642(1995).		
RL	[2]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=Plagemann;		
RC	MEDLINE=96010224; PubMed=7571421;		
RA	Faaberg K.S., Plagemann P.G.;		
RX	"The envelope proteins of lactate dehydrogenase-elevating virus and		
RT	their membrane topography."		
RT	Virology 212:512-525(1995).		
RL	[3]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=Plagemann;		
RC	STRAIN=Plagemann;		
RA	Palmer G.A.;		
RA	Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; U15146; AA085665.1; -.		
DR	PIR; JQ1997; JQ1997.		
DR	InterPro; IPR003434; PRRSV Env.		
DR	Pfam; PF02340; PRRSV Env; 1.		
DR	SEQUENCE 227 AA; 25998 MW; 16C8C9982DF2AAAF		
DR	CRC64;		

```

RESULT 37
Q9Y541
ID Q9Y541 PRELIMINARY; PRT; 227 AA.
AC Q9Y541;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Minor structural glycoprotein.
OS Lactate dehydrogenase-elevating virus (LDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
NCBI_TaxID=11048;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=LDV-v;
RC MEDLINE=99263186; PubMed=10329569;
RA Li K., Chen Z., Plagemann P.;
RT "High-frequency homologous genetic recombination of an arterivirus,
RT lactate dehydrogenase-elevating virus, in mice and evolution of
RT neuropathogenic variants.";
RL Virology 258:73-83(1999).
DR EMBL; AF092282; AAC84050.1; --
DR InterPro; IPR003434; PRRSV_Env.
DR Pfam; PF02340; PRRSV_Env; I.
SQ SEQUENCE 227 AA; 25844 MW; E2A5786358D9B813 CRC64;

Query Match 87.1%; Score 27; DB 12; Length 227;
Best Local Similarity 50.0%; Pred.No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

[illegible]

```

RESULT 39
Q9SL57
ID Q9SL57 PRELIMINARY; PRT; 239 AA.
AC Q9SL57;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE A2913070 protein.
GN A2913070.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC XCB1_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RP RC STRAIN=cv. Columbia;
RX RA MEDLINE=20083487; PubMed=10617197;
RX RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell L.R., Ketchum K.A., Lee J.-J., Rounsley C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768 (1999) .
RN [2]
SEQUENCE FROM N.A.
RP RC STRAIN=cv. Columbia;
RX RA

```


RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RA EMBL; AC006194; AAD25684.1; -;
 DR PIR; C84505; C84505.
 SQ SEQUENCE 239 AA; 26532 MW; 267943EF0F4AE25E CRC64;

Query Match 87.1%; Score 27; DB 10; Length 239;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXP 6
 Db 16 WAWAF 21

RESULT 40

ID Q95S28 PRELIMINARY; PRT; 244 AA.
 AC Q95S28;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE LD33016P.
 GN CG5427.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Merkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleb J., Paragass V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF060405; AAL25444.1; -;
 DR FlyBase; FBgn032433; CG5427.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR004156; OAPF_Nterm.
 DR Pfam; PF03132; OAPF_N; 1.
 SQ SEQUENCE 244 AA; 27696 MW; 901B73555CC239BA CRC64;

Query Match 87.1%; Score 27; DB 5; Length 244;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXP 6
 Db 95 WSWAF 100

RESULT 41

ID Q8X908 PRELIMINARY; PRT; 245 AA.
 AC Q8X908;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Orf, hypothetical protein.
 GN Z3983 OR EC3544.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaener J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta B.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22 (2001).
 DR EMBL; AE005497; AAG57791.1; -;
 DR EMBL; AP002562; BAB36967.1; -;
 DR PIR; C85916; C85916.
 DR PIR; H91071; H91071.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR004471; Azic.
 DR Pfam; PF03591; Azic; 1.
 KW Complete proteome.
 SQ SEQUENCE 245 AA; 26093 MW; 20AAF2D91E644DB3 CRC64;

Query Match 87.1%; Score 27; DB 16; Length 245;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WXXWXP 6
 Db 149 WSSWVF 154

RESULT 42

ID Q8FEQ6 PRELIMINARY; PRT; 245 AA.
 AC Q8FEQ6;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Hypothetical protein ygaZ.
 GN YGAZ OR C3235.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 DR EMBL; AE016765; BAN81687.1; -;
 DR PIR; C85916; C85916.
 DR PIR; H91071; H91071.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR004471; Azic.
 DR Pfam; PF03591; Azic; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 245 AA; 26093 MW; 20AAF2D91E644DB3 CRC64;

```
Query Match      87.1%; Score 27; DB 16; Length 245;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWVF 6
DB      149 WSSWVF 154

RESULT 43
Q83JZ6 PRELIMINARY; PRT; 245 AA.
ID Q83JZ6;
AC Q83JZ6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Orf, conserved hypothetical protein.
GN YGAZ OR SF2709 OR S2896.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RA MEDLINE=22272406; PubMed=12384590;
RX Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang J., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RA MEDLINE=22590274; PubMed=12704152;
RX Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL; AE015286; AA04202.1; -
DR EMBL; AE016987; AAP18029.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004471; AzlC.
DR Pfam; PF03591; AzlC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 245 AA; 26079 MW; 3BAC440092FE67F6 CRC64;

Query Match      87.1%; Score 27; DB 16; Length 245;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWVF 6
DB      149 WSSWVF 154

RESULT 44
Q829J1 PRELIMINARY; PRT; 247 AA.
ID Q829J1;
AC Q829J1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative lysozyme.
GN SAV6420.
OS Streptomyces avermitilis.

Query Match      87.1%; Score 27; DB 16; Length 245;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWVF 6
DB      149 WSSWVF 154

RESULT 45
Q8U324 PRELIMINARY; PRT; 255 AA.
ID Q8U324;
AC Q8U324;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PF0651.
GN PF0651.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010186; AAL80775.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 255 AA; 28359 MW; 51A21DF5FF37FC83 CRC64;

Query Match      87.1%; Score 27; DB 17; Length 255;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 WXXWVF 6
DB      137 WAIWTF 142
```

```

RESULT 46
Q8F276 PRELIMINARY; PRT; 270 AA.
ID Q8F276 AC
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Haloacid dehalogenase-like hydrolase.
GN LA2900.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56501 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011454; AAN50099.1; -.
GO GO:0016787; F:hydrolase activity; IEA.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 270 AA; 30695 MW; D765D449EE5C18D2 CRC64;

Query Match 87.1%; Score 27; DB 16; Length 270;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 WXXWKF 6
Db 102 WSSWIF 107

RESULT 47
Q8ZUD5 PRELIMINARY; PRT; 274 AA.
ID Q8ZUD5 AC
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Molybdopterin oxidoreductase, membrane subunit.
GN PA2836.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AB009896; AAL64472.1; -.
DR InterPro; IPR005614; NrfD.
DR Pfam; PF03916; NrfD; 1.
KW Complete proteome.
SQ SEQUENCE 274 AA; 28626 MW; B38FCFF487EE11A CRC64;

Query Match 87.1%; Score 27; DB 17; Length 274;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 WXXWKF 6
Db 268 WGAWAF 273

RESULT 48
Q89X05 PRELIMINARY; PRT; 275 AA.
ID Q89X05 AC
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BL10523 protein.
GN BL10523.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RL "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005936; BAC45788.1; -.
DR InterPro; IPR008934; AcPase_VanPerase.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPC; 1.
KW Complete proteome.
SQ SEQUENCE 275 AA; 30339 MW; 270F50635CF852E5 CRC64;

Query Match 87.1%; Score 27; DB 16; Length 275;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 WXXWKF 6
Db 76 WIAWAF 81

RESULT 49
O17009 PRELIMINARY; PRT; 293 AA.
ID O17009 AC
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T23B12.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Davidson S., Wohlmann P., Gillam B.;
RT "The sequence of C. elegans cosmid T23B12.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022982; AAB69940.1; -.
DR PIR; T32229; T32229.

```

```

DR WormPep; T23B12.10; CEI4044.
KW Hypothetical protein.
SQ SEQUENCE 293 AA; 32446 MW; 76646FCAACAACAC25 CRC64;

Query Match 87.1%; Score 27; DB 5; Length 293;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
44 WTIWTF 49

RESULT 50
O65422 PRELIMINARY; PRT; 294 AA.
ID AC O65422;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN F18E5.190 OR F17I22.30 OR AT4G21570.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
RA Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.-J., Voet M., Robben J.,
RA Volckaert G., Bancroft I., Mewes H.W., Mayer K.F.X., Schueller C.;
RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RA Van Der Schueren J., Vandenbussche F., Chuang Y.J., Braeken M.,
RA Robben J., Volckaert G., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RN SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [9]

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=A136;
RX MEDLINE=21225529; PubMed=11325934;
RA Kahng L.S., Shapiro L.;
RT "The Ccm DNA Methyltransferase of Agrobacterium tumefaciens Is
RT Essential, and Its Activity Is Cell Cycle Regulated.";
RL J. Bacteriol. 183:3085-3075(2001).
DR EMBL; AF327563; F:DNA binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002295; D2LN6_mtfase.
DR InterPro; IPR001091; Met trans CN4.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PRO0506; D2LN6MTFRASE.
DR PRINTS; PRO0508; S2LN4MTFRASE.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Methyltransferase; Transferase.
SQ SEQUENCE 381 AA; 42250 MW; 81150DCB5865C0F5 CRC64;

Query Match 92.1%; Score 35; DB 2; Length 381;
Best Local Similarity 66.7%; Pred. NO. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 352 WTFWHP 357

RESULT 12
Q8UH89 PRELIMINARY; PRT; 386 AA.
AC Q8UH89
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Adenine DNA methyltransferase.
GN ATU0794 OR AGR C 1453.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Zhang C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Yang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester B.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Okoroilo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmieu K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA William C., Allinger M., Doughty D., Scott C., Leppas C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;

```

```

RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AF009046; AAL41810.1; -
DR EMBL; AF008012; AAK86602.1; -
DR PIR; A97456; A97456.
DR PIR; AD2674; AD2674.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002295; D2LN6_mtfase.
DR InterPro; IPR001091; Met trans CN4.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PRO0506; D2LN6MTFRASE.
DR PRINTS; PRO0508; S2LN4MTFRASE.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Methyltransferase; Transferase; Complete proteome.
SQ SEQUENCE 386 AA; 42828 MW; 7717D611219F0965 CRC64;

Query Match 92.1%; Score 35; DB 16; Length 386;
Best Local Similarity 66.7%; Pred. NO. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 357 WTFWHP 362

RESULT 13
Q9AW80 PRELIMINARY; PRT; 394 AA.
AC Q9AW80
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative SAR DNA-binding protein-1.
GN NOPS.
OS Guillardia theta (Cryptomonas phi).
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20087226; PubMed=10618395;
RA Zauner S., Fraunholz M., Wastl J., Penny S.L., Beaton M.,
RA Cavalier-Smith T., Maier U., Douglas S.;
RT "Chloroplast protein and centrosomal genes, a tRNA intron, and odd
RT telomeres in an unusually compact eukaryotic genome, the cryptomonad
RT nucleomorph.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2123349; PubMed=11323671;
RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,
RA Wu X., Reith M., Cavalier-Smith T., Maier U.;
RT "The highly reduced genome of an enslaved algal nucleus.";
RL Nature 410:1091-1096(2001).
DR EMBL; AJ010592; CAC26989.1; -
DR PIR; D90105; D90105.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
KW DNA-binding.
SQ SEQUENCE 394 AA; 45642 MW; 98C3585C53E187C7 CRC64;

Query Match 92.1%; Score 35; DB 10; Length 394;
Best Local Similarity 66.7%; Pred. NO. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 WXXWHF 6
DB 169 WYXWHF 174

RESULT 14
Q8YFS6 PRELIMINARY; PRT; 403 AA.
ID Q8YFS6
AC Q8YFS6;
DT 01-WAR-2002 (TREMBlrel. 20, Created)
DT 01-WAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Adenine-specific methyltransferase (EC 2.1.1.72).
GN EMEI144.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16W / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kapatral V., Redkar R.J., Patra G., Mufier C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N., Overbeek R.;
RA "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
RM EMBL; AE009581; AAU52625.1; -.
DR PIR; AP3432; AF3432.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0009007; F:site-specific DNA-methyltransferase (adenin. .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002295; D2LN6_mtfase.
DR InterPro; IPR001091; Met_trans_CN4.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PR00506; D2LN6MTFRASE.
DR PRINTS; PR00508; S2LN4MTFRASE.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 403 AA; 45065 MW; 3ABB8BE232CB72F9 CRC64;

Query Match 92.1%; Score 35; DB 16; Length 403;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 373 WYFWHF 378

RESULT 15
Q7UNA5 PRELIMINARY; PRT; 419 AA.
ID Q7UNA5
AC Q7UNA5;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB7694.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Firellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RA "Complete genome sequence of the marine planctomycete Firellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
RM EMBL; BX294146; CAD75514.1; -.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 419 AA; 47048 MW; F5F68CB45747714D CRC64;

Query Match 92.1%; Score 35; DB 16; Length 419;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 174 WNAWHF 179

RESULT 16
Q9SBU4 PRELIMINARY; PRT; 421 AA.
ID Q9SBU4
AC Q9SBU4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Chloroplast w6 desaturase.
GN DRS6.
OS Chlamydomonas sp. (strain W80).
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=103365;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W80;
RA Miyasaka H.;
RA "Cloning of chloroplast w6 desaturase of a marine halotolerant
RT Chlamydomonas.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RM EMBL; AB031546; BAA83822.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase_1.
DR ProDom; PD001081; FA_desat_fam; 2.
SQ SEQUENCE 421 AA; 47932 MW; AD48F6039CBEDD57 CRC64;

Query Match 92.1%; Score 35; DB 10; Length 421;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 232 WAIWHF 237

RESULT 17
Q9ZRW0 PRELIMINARY; PRT; 454 AA.
ID Q9ZRW0
AC Q9ZRW0;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Nucleolar protein (Fragment).
OS Cicer arietinum (Chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OX NCBI_TaxID=3827;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=cv. castellana; TISSUE=Etolated epicotyl;
RA Duplico B., Esteban R., Labrador E.;
RT "A cDNA encoding a nucleolar protein is expressed in chickpea
epicotyls";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ012686; CAA10127.1; -.
DR InterPro: IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
FT NON_TER 1
SQ SEQUENCE 454 AA; 51134 MW; 26B173C03A8E2636 CRC64;

Query Match 92.1%; Score 35; DB 10; Length 454;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 124 WYSWHF 129

RESULT 18
Q7UJ09 PRELIMINARY; PRT; 469 AA.
AC Q7UJ09;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB11174.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Bozzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1";
RT Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
RL EMBL; BX294152; CAD77203.1; -.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 469 AA; 51193 MW; 02B4065710CB2A04 CRC64;

Query Match 92.1%; Score 35; DB 16; Length 469;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 333 WVAWHF 338

RESULT 19
O94514 PRELIMINARY; PRT; 497 AA.
ID O94514;
AC O94514;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Probable involvement in pre rRNA processing, putative sik1 protein
homolog.
DE homology.
GN SPC646.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Seeger K., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035216; CAA22814.1; -.
DR PIR; T40586; T40586.
DR GeneDB_Spombe; SPC646.10C; -.
DR InterPro: IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
SQ SEQUENCE 497 AA; 55388 MW; 27F94313F72AB987 CRC64;

Query Match 92.1%; Score 35; DB 3; Length 497;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 194 WYSWHF 199

RESULT 20
Q9LTV0 PRELIMINARY; PRT; 499 AA.
ID Q9LTV0;
AC Q9LTV0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Nucleolar protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
features of the regions of 4,504,864 bp covered by sixty P1 and PAC
clones";
RL DNA Res. 7:131-135 (2000).
DR EMBL; AB024033; BAB02430.1; -.
DR InterPro: IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
SQ SEQUENCE 499 AA; 56180 MW; 0800705886E821B8 CRC64;

Query Match 92.1%; Score 35; DB 10; Length 499;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 195 WYSWHF 200

RESULT 21
Q8XQC5 PRELIMINARY; PRT; 510 AA.
ID Q8XQC5;
AC Q8XQC5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative spermidine synthase transmembrane protein (EC 2.5.1.16).
GN SPC1 OR RSP1337 OR RS04762.
OS Ralstonia solanacearum (Pseudomonas solanacearum).

OG Plasmid megaplasmid.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae, Ralstonia.
 OX NCBI_TaxID=3305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 DR EMBL; AL646084; CAD18488.1; --
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
 DR GO; GO:0004766; F:spermidine synthase activity; IEA.
 DR GO; GO:0016740; E:transferase activity; IEA.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR000051; SAM_Bind.
 DR InterPro; IPR001045; Spermine synthase.
 DR Pfam; PF01564; Spermine synth. 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Transferase; Plasmid; Complete proteome.
 SQ SEQUENCE 510 AA; 56610 MW; 1814A88BDA2A56AB CRC64;
 Query Match 92.1%; Score 35; DB 16; Length 510;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 182 WTLWHF 187
 RESULT 22
 Q9SGT7 PRELIMINARY; PRT; 522 AA.
 ID Q9SGT7
 AC Q9SGT7
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Nucleolar protein (NOP56-like protein) (AT1G56110/T6H22_9) (SAR DNA
 binding protein, putative).
 GN T6H22.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Federpsiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
 RA Altafi H., Nguyen M., Lam B., Buehler E., Dunn P., Gonzalez A.,
 RA Kremetskaia I., Kim C., Lenz C., Li J., Liu S., Luros S.,
 RA Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S., Walker M., Yu G.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RP Phelan T.J., Spiker S.L.;
 RA "Arabidopsis MAR binding NOP56/58 homologs";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Tracy S.E., Ban J., Bowser L., Carninci P., Chung M.K.,
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,

RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
 RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Ban J., Bowser L.,
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis ORF clones";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 annotation";
 RL Genome Biol. 0:0-0(2002).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC009894; AAF02835.1; --
 DR EMBL; AF302492; AAG40838.1; --
 DR EMBL; AY039541; AAK62596.1; --
 DR EMBL; AY102151; AAM26718.1; --
 DR EMBL; AY087080; AAM64641.1; --
 DR FIR; D96602; D96602.
 DR InterPro; IPR002687; Nop.
 DR Pfam; PF01798; Nop; 1.
 DR ProDom; PD004104; Nop; 1.
 SQ SEQUENCE 522 AA; 58673 MW; 00C2FA0BF7C23FBB CRC64;
 Query Match 92.1%; Score 35; DB 10; Length 522;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 195 WYSWHF 200
 RESULT 23
 Q8XQC8 PRELIMINARY; PRT; 525 AA.
 ID Q8XQC8
 AC Q8XQC8;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Putative spermidine synthase protein (EC 2.5.1.16).
 GN SPE2 OR RSP1306 OR RS05693.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=3305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissbach J., Boucher C.A.;


```

RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002);
RL EMBL: AL646034; CAD18457.1; -.
DR GO: 0046821; C:extrachromosomal DNA; IEA.
DR GO: 0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR GO: 0004766; F:sermidine synthase activity; IEA.
DR GO: 0016740; F:transferase activity; IEA.
DR InterPro: IPR000051; SAM bind.
DR InterPro: IPR001045; Srimine synthase.
DR Pfam: PF01564; Spermine synth; 1.
KW Transferase; Plasmid; Complete proteome.
SQ SEQUENCE 525 AA; 58094 MW; 60ACBEDBA8CC3B34 CRC64;

Query Match 92.1%; Score 35; DB 16; Length 525;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 197 WTLWHF 202

RESULT 24
Q8H2U5 PRELIMINARY; PRT; 549 AA.
AC Q8H2U5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative nucleolar protein.
GN P0453B03.11.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone:P0453B03.1";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005452; BAC22425.1; -.
DR InterPro: IPR002687; Nop.
DR Pfam: PF01798; Nop; 1.
DR ProDom: PD004104; Nop; 1.
SQ SEQUENCE 549 AA; 61307 MW; 4A61D5AFFB5E5E5AB CRC64;

Query Match 92.1%; Score 35; DB 10; Length 549;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 206 WYSWHF 211

RESULT 25
Q811I3 PRELIMINARY; PRT; 594 AA.
AC Q811I3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN Pf11.0191.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;

RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA Partea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
falciparum.";
RL Nature 419:498-511(2002).
DR EMBL: AE014838; AAN35775.1; -.
DR InterPro: IPR002687; Nop.
DR Pfam: PF01798; Nop; 1.
DR ProDom: PD004104; Nop; 1.
KW Hypothetical protein.
SQ SEQUENCE 594 AA; 69389 MW; 758D3CD63902EB34 CRC64;

Query Match 92.1%; Score 35; DB 5; Length 594;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 193 WYSWHF 198

RESULT 26
Q9CLS4 PRELIMINARY; PRT; 187 AA.
AC Q9CLS4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PM1135.
GN PM1135.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70.
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL: AE006154; AAK03219.1; -.
DR GO: 0003700; F:transcription factor activity; IEA.
DR GO: 0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001647; HTH_Tetr.
DR Pfam: PF00440; tetr; 1.
DR PRINTS: PR00455; HTHTETR.
KW DNA-binding; Hypothetical protein; Transcription regulation;
KW Complete proteome.
SQ SEQUENCE 187 AA; 22325 MW; 28E3E9B5C14115F CRC64;

Query Match 89.5%; Score 34; DB 16; Length 187;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 89 WNIWHF 94

RESULT 27
Q96N05 PRELIMINARY; PRT; 190 AA.
ID Q96N05

```

```

Q9GN05;
AC 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ31609.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056171; BAB71110.1; -
KW Hypothetical protein.
SQ SEQUENCE 190 AA; 21014 MW; B43B5CBE1B14627F CRC64;

Query Match 89.5%; Score 34; DB 4; Length 190;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 82 WGRWHF 87

RESULT 28
Q9C600
ID Q9C600 PRELIMINARY; PRT; 201 AA.
AC Q9C600;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (fragment).
GN FAD6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Fourmann M., Froger N., Brunel D.;
RA "Amplified consensus gene markers: Tools designing for a genetic map
of Arabidopsis-known-function genes in Brassica.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229392; AAK00664.1; -
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase_1.
DR ProDom; PD001081; FA_desat_fam; 1.
FT NON_TER 1
FT NON_TER 201
SQ SEQUENCE 201 AA; 23084 MW; 9261EFC0BD3771C0 CRC64;

Query Match 89.5%; Score 34; DB 10; Length 201;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 156 WVNWHF 161

```

```

RESULT 29
Q9AXY3
ID Q9AXY3 PRELIMINARY; PRT; 202 AA.
AC Q9AXY3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (fragment).
GN FAD6-BR-1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RA Fourmann M., Froger N., Brunel D.;
RA "Amplified consensus gene markers: Tools designing for a genetic map
of Arabidopsis-known-function genes in Brassica.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229390; AAK00662.1; -
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase_1.
DR ProDom; PD001081; FA_desat_fam; 1.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23163 MW; A96E403AAB1EC008 CRC64;

Query Match 89.5%; Score 34; DB 10; Length 202;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 157 WVNWHF 162

RESULT 30
Q9AXY5
ID Q9AXY5 PRELIMINARY; PRT; 213 AA.
AC Q9AXY5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (fragment).
GN FAD6-BN-2.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RA Fourmann M., Froger N., Brunel D.;
RA "Amplified consensus gene markers: Tools designing for a genetic map
of Arabidopsis-known-function genes in Brassica.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229388; AAK00660.1; -
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase_1.
DR ProDom; PD001081; FA_desat_fam; 2.
FT NON_TER 1
FT NON_TER 213
SQ SEQUENCE 213 AA; 24771 MW; D30EB4E98ADBDC06 CRC64;

Query Match 89.5%; Score 34; DB 10; Length 213;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 WXXWHF 6
Db 154 WVNWHF 159

RESULT 31
Q9AXY2 PRELIMINARY; PRT; 216 AA.
AC Q9AXY2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (Fragment).
GN FAD6-BN-1.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Stellar;
RA Fourmann M., Froger N., Brunel D.; Tools designing for a genetic map
RT of Arabidopsis-known-function genes in Brassica.
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229391; AAK00663.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase; 1.
DR ProDom; PD001081; FA_desat_fam; 2.
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 25110 MW; EFFCEDBB1B0E3B4 CRC64;

Query Match
Best Local Similarity 89.5%; Score 34; DB 10; Length 216;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 157 WVNWHF 162

RESULT 32
Q9AXY4 PRELIMINARY; PRT; 222 AA.
AC Q9AXY4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Omega 6 reductase (Fragment).
GN FAD6-BO-1.
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Rapid Cycling;
RA Fourmann M., Froger N., Brunel D.; Tools designing for a genetic map
RT of Arabidopsis-known-function genes in Brassica.
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229389; AAK00661.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase; 1.
DR ProDom; PD001081; FA_desat_fam; 2.
FT NON_TER 1
FT NON_TER 222
SQ SEQUENCE 222 AA; 25761 MW; A2B19F83893BCD4E CRC64;

Query Match
Best Local Similarity 89.5%; Score 34; DB 16; Length 230;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 26 WVNWHF 31

RESULT 34
Q7UZY3 PRELIMINARY; PRT; 230 AA.
AC Q7UZY3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

```

Query Match
Best Local Similarity 89.5%; Score 34; DB 10; Length 222;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 163 WVNWHF 169

RESULT 33
Q53604 PRELIMINARY; PRT; 230 AA.
AC Q53604;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Rv0059.
GN Rv0059 OR MT030.02 OR MT0065.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL021428; CAA16240.1; -.
DR EMBL; AB006919; AAK44287.1; -.
DR PIR; D70847; D70847.
DR TIGR; MT0065; -.
DR TubercuList; Rv0059; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 230 AA; 25580 MW; 1A109D34B16590A8 CRC64;

Query Match
Best Local Similarity 89.5%; Score 34; DB 16; Length 230;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 26 WVNWHF 31

RESULT 34
Q7UZY3 PRELIMINARY; PRT; 230 AA.
AC Q7UZY3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

```

DE Hypothetical protein.
GN ME0060.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=AF2122/97;
RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempé C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
RL EMBL; BX248334; CAD92922.1; -.
KW Complete proteome.
SQ SEQUENCE 230 AA; 25580 MW; 1A109D34B16590A8 CRC64;

Query Match 89.5%; Score 34; DB 16; Length 230;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 26 WIVWHF 31

RESULT 35
Q8EWI8 PRELIMINARY; PRT; 297 AA.
AC Q8EWI8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN MYE2150.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL; AP004170; BAC44006.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 297 AA; 34216 MW; 620C276293698E8B CRC64;

Query Match 89.5%; Score 34; DB 16; Length 297;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 119 WLLWHF 124

RESULT 36
Q9EXF1 PRELIMINARY; PRT; 310 AA.
AC Q9EXF1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Putative phosphotransferase D-arabitol specific component IIC
DE (Fragment).
GN ATLC.

```

```

OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=LO28;
RX MEDLINE=21212979; PubMed=11316371;
RA Saklani-Jusforgues H., Fontan B., Goossens P.L.;
RT "Characterisation of a Listeria monocytogenes mutant deficient in D-
RT arabinol fermentation.";
RL Res. Microbiol. 152:175-177(2001).
DR EMBL; AJ292552; CAC20641.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR InterPro; IPR004703; Gal_spec_IIC.
DR Pfam; PF03611; EIIC-GAT; 1.
KW Transferase.
FT NON_TER 1 310
FT TER 310
SQ SEQUENCE 310 AA; 33711 MW; A2F9F89C8C8DACF CRC64;

Query Match 89.5%; Score 34; DB 2; Length 310;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 18 WNYWHF 23

RESULT 37
Q8RD64 PRELIMINARY; PRT; 415 AA.
AC Q8RD64;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein TTE0183.
GN TTE0183.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AF012991; AAM23484.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR InterPro; IPR004703; Gal_spec_IIC.
DR Pfam; PF03611; EIIC-GAT; 1.
DR TIGRFAMs; TIGR00827; EIIC-GAT; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 415 AA; 44078 MW; 6FFCD2BAAC6C29DD CRC64;

Query Match 89.5%; Score 34; DB 16; Length 415;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 122 WNYWHF 127

RESULT 38
Q8H0T7

```

```

ID Q8H0T7 PRELIMINARY; PRT; 423 AA.
AC Q8H0T7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SAR DNA-binding protein-like (Fragment).
GN A75G27120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Garninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT02060; AAN72071.1; -
DR GO; GO:0003677; F-DNA binding; IEA.
DR InterPro; IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
KW DNA-binding.
FT NON TER 423 423
SQ SEQUENCE 423 AA; 46318 MW; 912587CAAD773750 CRC64;

Query Match 89.5%; Score 34; DB 10; Length 423;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 184 WFGWHP 189

RESULT 39
Q927H4 PRELIMINARY; PRT; 423 AA.
ID Q927H4;
AC Q927H4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein lin2814.
GN LIN2814.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CLIP 11262 / Serovar 6a;
RC MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madieno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordieck G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RL "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596173; CAC98040.1; -
DR PIR; AH1783; AH1783.
DR ListList; LIN02814; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009401; P-phosphoenolpyruvate-dependent sugar phospho. . .; IEA.

```

```

DR InterPro; IPR004703; Gal_spec IIC.
DR Pfam; PF03611; EIIC-GAT; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 423 AA; 45557 MW; CB548E19211EC541 CRC64;

Query Match 89.5%; Score 34; DB 16; Length 423;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 126 WNYWHF 131

RESULT 40
Q8Y412 PRELIMINARY; PRT; 423 AA.
ID Q8Y412;
AC Q8Y412;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein lmo2665 (Galactitol-specific PTS enzyme IIC component).
GN LMO2665.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madieno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordieck G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RL "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591984; CAD00878.1; -
DR PIR; AH1407; AH1407.
DR ListList; LMO02665; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009401; P-phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR InterPro; IPR004703; Gal_spec_IIC.
DR Pfam; PF03611; EIIC-GAT; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 423 AA; 45572 MW; 885B73CE43C8BEA5 CRC64;

Query Match 89.5%; Score 34; DB 16; Length 423;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 126 WNYWHF 131

RESULT 41
O48663

```

ID O48663 PRELIMINARY; PRT; 424 AA.
 AC O48663;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Chloroplast w6 desaturase.
 GN DES6.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98158334; PubMed=9498569;
 RA Sato N., Fujiwara S., Kawaguchi A., Tsuzuki M.;
 RT "Cloning of a gene for chloroplast w6 desaturase of a green alga,
 RT Chlamydomonas reinhardtii.";
 RL J. Biochem. 122:1224-1232(1997).
 DR EMBL; AB007640; BAA3881.1; -;
 DR PIR; JC5891.
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR InterPro; IPR005804; FA desat fam.
 DR Pfam; PF00487; FA desaturase; 1;
 DR ProDom; PD001081; FA desat_fam; 2;
 SQ SEQUENCE 424 AA; 48377 MW; D553054CICEB6B0D CRC64;

 Query Match 89.5%; Score 34; DB 10; Length 424;
 Best Local Similarity 66.7%; Pred. No. 3.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 Qy 1 WXXWHF 6
 Db 234 WLVWHF 239

 RESULT 42
 ID O8ZL19 PRELIMINARY; PRT; 452 AA.
 AC O8ZL19;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative PTS system galactitol-specific enzyme IIC component.
 GN STM3782
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 DR EMBL; AE008876; AAL22640.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
 DR InterPro; IPR004703; Gal_spec_IIC.
 DR Pfam; PF03611; EIIC-GAT; 1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 452 AA; 48561 MW; C29CE70D3B6B4D39 CRC64;

 Query Match 89.5%; Score 34; DB 16; Length 452;
 Best Local Similarity 66.7%; Pred. No. 3.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 Qy 1 WXXWHF 6
 Db 1 WXXWHF 6

Db 130 WNYWHF 135

 RESULT 43
 ID O8Z2K1 PRELIMINARY; PRT; 452 AA.
 AC O8Z2K1;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative PTS system IIC component (Putative IIC component of PTS
 DE system).
 GN STY4000 OR T3736.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth I., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 DR EMBL; AL627280; CAD03212.1; -;
 DR GO; GO:0016846; AO071229.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
 DR InterPro; IPR004703; Gal_spec_IIC.
 DR Pfam; PF03611; EIIC-GAT; 1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 452 AA; 48561 MW; D4E65BCB41D7859A CRC64;

 Query Match 89.5%; Score 34; DB 16; Length 452;
 Best Local Similarity 66.7%; Pred. No. 3.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 Qy 1 WXXWHF 6
 Db 130 WNYWHF 135

 RESULT 44
 ID O8X5T1 PRELIMINARY; PRT; 462 AA.
 AC O8X5T1;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative phosphotransferase system enzyme subunit (Putative
 DE phosphotransferase system enzyme IIC).
 GN Z4877 OR EC54352.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.

```

OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=0157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RT Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=0157:H7 / RMD 050952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:111-22 (2001).
DR EMBL; AE005572; AAG58612.1; -.
DR EMBL; AF025655; BAB3775.1; ALT_INIT.
DR PIR; H86018; H86018.
DR PIR; H91172; H91172.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR InterPro; IPR004703; Gal_spec_IIC.
DR Pfam; PF03611; EIIC-GAT; 1.
DR Transferase; Complete proteome.
KW SEQUENCE 462 AA; 50039 MW; FA02CF54A037BA66 CRC64;
SQ
Query Match 89.5%; Score 34; DB 16; Length 462;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 WXXWHF 6
DB 140 WNYWHF 145
RESULT 45
Q8FCM4 PRELIMINARY; PRT; 462 AA.
ID Q8FCM4;
AC Q8FCM4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-JUN-2003 (TREMBLrel. 23, Last sequence update)
DE PTS system, galactitol-specific IIC component.
GN C4279.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016768; AAN82715.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR InterPro; IPR004703; Gal_spec_IIC.
DR Pfam; PF03611; EIIC-GAT; 1.
DR Transferase; Complete proteome.
KW SEQUENCE 465 AA; 49818 MW; F652B568A443D3D06 CRC64;
SQ
Query Match 89.5%; Score 34; DB 16; Length 465;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 WXXWHF 6
DB 123 WNYWHF 128
RESULT 47
Q8IJV7 PRELIMINARY; PRT; 469 AA.
ID Q8IJV7;
AC Q8IJV7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-JUN-2003 (TREMBLrel. 23, Last sequence update)
DE Nucleolar protein NOP5, putative.
GN PF10.0085.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,

```

```

RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrrell B.,
RA "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014830; AAN35283.1; -.
DR InterPro; IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
SQ SEQUENCE 469 AA; 53410 MW; D66822EAF6AD0FA CRC64;

Query Match 89.5%; Score 34; DB 5; Length 469;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 189 WYQWHF 194

RESULT 48
ID Q8YMN8 PRELIMINARY; PRT; 472 AA.
AC Q8YMN8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein A114895.
GN ALI4895.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT Cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003598; BAB76594.1; -.
DR FIR; AG2417; AG2417.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR004294; RPE65.
DR InterPro; IPR001680; WD40.
DR Pfam; PF03055; RPE65; 1.
DR PROSITE; PS00678; WD REPEATS 1; 1.
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 472 AA; 52993 MW; E78632608C1B5616 CRC64;

Query Match 89.5%; Score 34; DB 16; Length 472;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 285 WYQWHF 290

RESULT 49
ID Q9BHU7 PRELIMINARY; PRT; 473 AA.
AC Q9BHU7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein A114895.
GN ALI4895.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT Cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003598; BAB76594.1; -.
DR FIR; AG2417; AG2417.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR004294; RPE65.
DR InterPro; IPR001680; WD40.
DR Pfam; PF03055; RPE65; 1.
DR PROSITE; PS00678; WD REPEATS 1; 1.
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 472 AA; 52993 MW; E78632608C1B5616 CRC64;

Query Match 89.5%; Score 34; DB 11; Length 473;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 197 WYQWHF 202

RESULT 50
ID Q70396 PRELIMINARY; PRT; 473 AA.
AC Q70396;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SIK similar protein.
GN NOL5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=C3H/HeJ;
RA Zierke M., Martin M.U.;
RT "cDNA for mouse SIK similar protein (MSSP).";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF053232; AAC08435.1; -.
DR MGD; MGI:1933184; NOL5.
DR DR GO; GO:0005730; C:nucleolus; ISS.
DR DR GO; GO:0003754; F:chaperone activity; NAS.
DR DR GO; GO:0030515; F:snRNA binding; ISS.
DR DR GO; GO:0016049; P:cell growth; ISS.
DR DR GO; GO:0006364; P:rRNA processing; ISS.
DR InterPro; IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
SQ SEQUENCE 473 AA; 53111 MW; 6C7678A2B87D3023 CRC64;

Query Match 89.5%; Score 34; DB 11; Length 473;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable nucleolar protein involved in pre-rRNA processing.
GN P1046.54.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=Friedlin;
RA Aert R., Robben J., Volckaert G., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL359683; CAC37159.2; -.
DR InterPro; IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
SQ SEQUENCE 473 AA; 52685 MW; E78CFF8667F28F54 CRC64;

Query Match 89.5%; Score 34; DB 5; Length 473;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 197 WYQWHF 202

RESULT 50
ID Q70396 PRELIMINARY; PRT; 473 AA.
AC Q70396;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SIK similar protein.
GN NOL5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=C3H/HeJ;
RA Zierke M., Martin M.U.;
RT "cDNA for mouse SIK similar protein (MSSP).";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF053232; AAC08435.1; -.
DR MGD; MGI:1933184; NOL5.
DR DR GO; GO:0005730; C:nucleolus; ISS.
DR DR GO; GO:0003754; F:chaperone activity; NAS.
DR DR GO; GO:0030515; F:snRNA binding; ISS.
DR DR GO; GO:0016049; P:cell growth; ISS.
DR DR GO; GO:0006364; P:rRNA processing; ISS.
DR InterPro; IPR002687; Nop.
DR Pfam; PF01798; Nop; 1.
DR ProDom; PD004104; Nop; 1.
SQ SEQUENCE 473 AA; 53111 MW; 6C7678A2B87D3023 CRC64;

Query Match 89.5%; Score 34; DB 11; Length 473;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6

```


Db 123 WYGHF 128

Search completed: June 10, 2004, 10:50:55
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:51:30 ; Search time 46.3333 Seconds

(without alignments)
36.589 Million cell updates/sec

Title: US-09-912-414-2

Perfect score: 45

Sequence: 1 WVRWHF 6

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1596107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 447529

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	100.0	6	3	AAB01492	Aab01492 Peptide w
2	41	91.1	6	3	AAB01505	Aab01505 Peptide w
3	39	86.7	6	3	AAB01509	Aab01509 Peptide w
4	39	86.7	6	3	AAB01506	Aab01506 Peptide w
5	35	77.8	6	3	AAB01508	Aab01508 Peptide w
6	34	75.6	6	3	AAB01504	Aab01504 Peptide w
7	34	75.6	6	3	AAB01499	Aab01499 Peptide w
8	34	75.6	8	2	AAR60429	Aar60429 Antiprol
9	34	75.6	8	2	AAR60444	Aar60444 Antiprol
10	31	68.9	6	2	AAR37389	Aar37389 Peptide f
11	31	68.9	6	2	AAW28912	Aaw28912 Opioid pe
12	31	68.9	6	2	AAR93770	Aar93770 New pepti
13	31	68.9	6	2	AAW23019	Aaw23019 Opioid pe
14	31	68.9	6	3	AAB01507	Aab01507 Peptide w
15	31	68.9	7	4	AAW45777	Aaw45777 H11 bindi
16	31	68.9	10	2	AAR86140	Aar86140 Anti-ELAM
17	31	68.9	10	2	AAR86145	Aar86145 Anti-ELAM
18	31	68.9	10	2	AAR86146	Aar86146 Anti-ELAM
19	31	68.9	10	2	AAW63363	Aaw63363 ELAM-1 pe
20	31	68.9	10	2	AAW63364	Aaw63364 ELAM-1 pe
21	31	68.9	10	2	AAW63358	Aaw63358 ELAM-1 pe
22	30	66.7	6	2	AAR37390	Aar37390 Peptide f
23	30	66.7	6	2	AAW66066	Aaw66066 Peptide u
24	30	66.7	6	2	AAW24292	Aaw24292 Somatosta
25	30	66.7	6	6	ABR45592	AbR45592 Staphyloc

26	30	66.7	8	2	AAW08189	Aay08189 Clotting
27	30	66.7	12	2	AAW80380	Aaw80380 Peptide e
28	30	66.7	12	6	ABR75359	ABr75359 Biologica
29	30	66.7	14	5	ABB74383	Abb74383 Karyophil
30	29	64.4	7	4	AAB49729	Aab49729 Peptide S
31	29	64.4	8	5	ABB90493	Abb90493 Homidae
32	28	62.2	6	2	AAR33522	Aar33522 Peptide f
33	28	62.2	6	2	AAR37388	Aar37388 Peptide f
34	28	62.2	6	2	AAR93706	Aar93706 Cyclo-Ty
35	28	62.2	6	2	AAR93707	Aar93707 Cyclo-Ty
36	28	62.2	6	2	AAR93709	Aar93709 Cyclo-Ty
37	28	62.2	6	2	AAR93719	Aar93719 Cyclo-Ty
38	28	62.2	10	2	AAR74033	Aar74033 Bombesin-
39	28	62.2	10	2	AAR86144	Aar86144 Anti-ELAM
40	28	62.2	10	2	AAW63962	Aaw63962 ELAM-1 pe
41	28	62.2	12	2	AAR36519	Aar36519 D32.39 an
42	28	62.2	12	2	AAR56756	Aar56756 Random pe
43	28	62.2	12	2	AAR91504	Aar91504 D32.39 mo
44	28	62.2	12	2	AAW25286	Aaw25286 Antibody
45	28	62.2	12	4	AAW60032	Aab60032 Internali

ALIGNMENTS

RESULT 1
AAB01492
ID AAB01492 standard; peptide; 6 AA.
XX
AC AAB01492;
XX
XX 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
XX WPI; 2000-532806/48.
XX
PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Claim 6; Page 2; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
SQ Sequence 6 AA;
Query Match 100.0%; Score 45; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWHF 6
Db |||||
1 WVRWHF 6

RESULT 2
AAB01505
ID AAB01505 standard; peptide; 6 AA.
XX
AC AAB01505;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Example; Page 26; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
SQ Sequence 6 AA;
Query Match 91.1%; Score 41; DB 3; Length 6;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWHF 6
Db 1 WVRWHF 6

RESULT 3
AAB01509
ID AAB01509 standard; peptide; 6 AA.
XX
AC AAB01509;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
```

```
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Example; Page 26; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
SQ Sequence 6 AA;
Query Match 86.7%; Score 39; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WVRWH 5
Db 1 WVRWH 5

RESULT 4
AAB01506
ID AAB01506 standard; peptide; 6 AA.
XX
AC AAB01506;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
```

XX Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Example; Page 26; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
XX Sequence 6 AA;
SQ

Query Match 86.7%; Score 39; DB 3; Length 6;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWHF 6
DB 1 WVRWHF 6

RESULT 5
AAB01508
ID AAB01508 standard; peptide; 6 AA.
XX
AC AAB01508;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Example; Page 26; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
XX Sequence 6 AA;
SQ

Query Match 77.8%; Score 35; DB 3; Length 6;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWHF 6
DB 1 WVRWHF 6

RESULT 5
AAB01508
ID AAB01508 standard; peptide; 6 AA.
XX
AC AAB01508;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Example; Page 26; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
XX Sequence 6 AA;
SQ

Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WVRWHF 6
DB 1 WVRWHF 6

RESULT 6
AAB01504
ID AAB01504 standard; peptide; 6 AA.
XX
AC AAB01504;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
OS Synthetic.
XX
PN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Example; Page 26; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
XX Sequence 6 AA;
SQ

Query Match 75.6%; Score 34; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 VRWHF 6
DB 2 VRWHF 6

RESULT 7
AAB01499
ID AAB01499 standard; peptide; 6 AA.
XX
AC AAB01499;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX

KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
 KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
 KW restenosis.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 2 /note= "Any amino acid"
 FT Misc-difference 3 /note= "Any amino acid"
 FT
 FT
 XX WO200044771-A1.
 XX
 XX PD 03-AUG-2000.
 XX
 XX PF 26-JAN-2000; 2000WO-GB000227.
 XX
 XX PR 26-JAN-1999; 99GB-00001710.
 XX
 XX (PROL-) PROLIFIX LTD.
 XX
 XX PI Mueller R, Kontermann RE, Montigiani S;
 XX
 XX DR WPI; 2000-532806/48.
 XX
 XX PT Peptides binding to the DNA binding domain of transcription factor E2F
 XX and inhibiting cell cycle progression, useful for the treatment of
 XX cancer.
 XX
 XX PS Claim 4; Page 9; 42pp; English.
 XX
 XX CC Peptides which bind to the DNA binding domain of transcription factor E2F
 XX and inhibit cell cycle progression may be useful as research agents to
 XX investigate the interaction between E2F and DP-1, or the activation of
 XX transcription by E2F-1/DP-1 heterodimers. They may also be used for
 XX inducing apoptosis and/or cell cycle arrest in a cell, particularly for
 XX treatment of cancer or other proliferative disorders such as psoriasis
 XX and restenosis
 XX
 XX SQ Sequence 6 AA;
 Query Match 75.6%; Score 34; DB 3; Length 6;
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WVRWHF 6
 Db 1 WXXWHF 6
 RESULT 8
 AAR60429
 ID AAR60429 standard; peptide; 8 AA.
 XX
 XX AC AAR60429;
 XX
 XX DT 25-MAR-2003 (revised)
 XX DT 30-MAR-1995 (first entry)
 XX
 XX DE Antiproliferative peptide to transplantable human B-cell lymphoma.
 XX
 XX KW antiproliferative; transplant; B-cell lymphoma line SUP-B8; Burkitt's;
 KW inhibit clonal expansion; induce apoptosis; anti-idiotype; IgM lambda;
 KW inhibit cell proliferation; peptidomimetics; cell surface receptor;
 KW immunoglobulin superfamily; treatment; neoplasia; identification;
 KW induce replication; therapy; clonal anergy; modulate tyrosine kinase.
 XX
 XX OS Synthetic.
 XX
 XX FN WO9418345-A1.
 XX PD 18-AUG-1994.
 XX
 XX PF 04-FEB-1994; 94WO-US001319.
 XX
 XX PR 05-FEB-1993; 93US-00014426.
 XX PD 18-AUG-1994.

XX
 PF 04-FEB-1994; 94WO-US001319.
 XX
 PR 05-FEB-1993; 93US-00014426.
 PR 15-NOV-1993; 93US-00153341.
 XX
 XX (AFFY-) AFFYMAX TECHNOLOGIES NV.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 XX PI Renschler ME, Levy R, Bhatt RR, Dower WJ;
 XX
 XX DR WPI; 1994-279762/34.
 XX
 XX PT Identifying anti-proliferative peptide(s) which specifically bind to
 XX immunoglobulin super-family species idiotype - esp. to inhibit B-cell
 XX lymphoma and leukocytic leukaemia cell proliferation, for anti-idiotype
 XX therapy.
 XX
 XX PS Claim 7; Page 45; 69pp; English.
 XX
 XX CC AAR60400-73 are peptide ligands which bind to purified IgM lambda
 XX receptor of the human Burkitt's lymphoma cell line SUP-B8. Peptides
 XX AAR60414 to AAR60473 were biotinylated and linked to streptavidin. The
 XX peptides were identified with the use of filamentous phage libraries
 XX displaying random peptides. Corresponding synthetic peptides bound
 XX specifically to this Ig receptor, and blocked the binding of an anti-
 XX idiotype antibody. The ligands, when conjugated to form dimers or
 XX tetramers, induced cell death by apoptosis in vitro at nanomolar
 XX concentrations. This effect was associated with the specific stimulation
 XX of intracellular protein tyrosine phosphorylation. The peptides of the
 XX invention can be used individually, as complexes of cross-linked peptides
 XX or can be conjugated to deliver toxins or radionuclides to neoplastic
 XX cells bearing the specific Ig receptor. (Updated on 25-MAR-2003 to
 XX correct PN field.)
 XX
 XX SQ Sequence 8 AA;
 Query Match 75.6%; Score 34; DB 2; Length 8;
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db 3 WVRWH 7
 RESULT 9
 AAR60444
 ID AAR60444 standard; peptide; 8 AA.
 XX
 XX AC AAR60444;
 XX
 XX DT 25-MAR-2003 (revised)
 XX DT 30-MAR-1995 (first entry)
 XX
 XX DE Antiproliferative peptide to transplantable human B-cell lymphoma.
 XX
 XX KW antiproliferative; transplant; B-cell lymphoma line SUP-B8; Burkitt's;
 KW inhibit clonal expansion; induce apoptosis; anti-idiotype; IgM lambda;
 KW inhibit cell proliferation; peptidomimetics; cell surface receptor;
 KW immunoglobulin superfamily; treatment; neoplasia; identification;
 KW induce replication; therapy; clonal anergy; modulate tyrosine kinase.
 XX
 XX OS Synthetic.
 XX
 XX FN WO9418345-A1.
 XX PD 18-AUG-1994.
 XX
 XX PF 04-FEB-1994; 94WO-US001319.
 XX
 XX PR 05-FEB-1993; 93US-00014426.
 XX PR 15-NOV-1993; 93US-00153341.

XX (AFFY-) AFFYMAX TECHNOLOGIES NV.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Renschler MF, Levy R, Bhatt RR, Dower WJ;
 XX WPI; 1994-279762/34.
 XX
 XX Identifying anti-proliferative peptide(s) which specifically bind to
 PT immunoglobulin super-family species idioType - esp. to inhibit B-cell
 PT lymphoma and leukocytic leukaemia cell proliferation, for anti-idiotypic
 PT therapy.
 XX
 PS Claim 7; Page 45; 69pp; English.
 XX
 CC AAR60400-73 are peptide ligands which bind to purified IgM lambda
 CC receptor of the human Burkitt's lymphoma cell line SUP-B8. Peptides
 CC AAR60414 to AAR60473 were biotinylated and linked to streptavidin. The
 CC peptides were identified with the use of filamentous phage libraries
 CC displaying random peptides. Corresponding synthetic peptides bound
 CC specifically to this Ig receptor, and blocked the binding of an anti-
 CC idioType antibody. The ligands, when conjugated to form dimers or
 CC tetramers, induced cell death by apoptosis in vitro at nanomolar
 CC concentrations. This effect was associated with the specific stimulation
 CC of intracellular protein tyrosine phosphorylation. The peptides of the
 CC invention can be used individually, as complexes of cross-linked peptides
 CC or can be conjugated to deliver toxins or radionuclides to neoplastic
 CC cells bearing the specific Ig receptor. (Updated on 25-MAR-2003 to
 CC correct FN field.)
 XX
 SQ Sequence 8 AA;
 Query Match 75.6%; Score 34; DB 2; Length 8;
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WVRWH 5
 Db | | | |
 3 WYRWH 7
 RESULT 10
 ID AAR37389
 AC AAR37389 standard; peptide; 6 AA.
 AC AAR37389;
 DT 07-JUL-1993 (first entry)
 DE Peptide for treating septic shock.
 DE Toxic shock; blood endotoxin removal; serum; diagnostic reagent;
 KW cytokine release control; treatment; pertussis; bacterial meningitis;
 KW HIV related infections; polymyxin B; Group II.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 PH Region 1..3 /note= "repeat region"
 FT Region 4..6 /note= "repeat region"
 FT
 XX ZA9200943-A.
 XX 25-NOV-1992.
 PD 10-FEB-1992; 92ZA-00000943.
 PF 11-FEB-1991; 91US-00658744.
 PR (PORR/) PORRO M.
 PA
 XX

PI Porro M;
 XX WPI; 1993-094304/11.
 XX
 DR New peptide for treatment or prevention of toxic shock - comprises
 XX specified sequences of aminoacid(s) and analogs comprising sequences
 PT retro-orientated.
 PT
 XX Example; Page 5; 39pp; English.
 XX
 CC The (Group II) peptide is an example of a generic peptide of formula R-(
 CC Lys/Arg/His - Phe/Tyr/Trip - Leu/Ile/Val)n-R, where n = 1-100 and each R
 CC is H, an amino acid residue or a fatty acid residue. The peptide is
 CC useful for treating or preventing septic shock, mixing with polymyxin B
 CC to reduce its toxicity; removing endotoxins from blood, sera or other
 CC fluids (in vivo or in vitro); controlling release of cytokines induced by
 CC endotoxins; as diagnostic reagents to detect and quantify toxins in blood
 CC or sera; preparing non-toxic antigenic complexes of lipid A or
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial
 CC meningitis and HIV-related infections. The usual dose is 10-100
 CC ug/kg/day, given parenterally. It binds to the same sites as polymyxin B,
 CC i.e. it inhibits all the toxic effects of lipid A. It has no antibiotic
 CC activity; does not lyse erythrocytes; has no toxicity in mice when
 CC injected at 50mg/kg and is relatively unstable against proteases
 XX
 SQ Sequence 6 AA;
 Query Match 68.9%; Score 31; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVRW 4
 Db | | | |
 2 WVRW 5
 RESULT 11
 AAW28912
 ID AAW28912 standard; peptide; 6 AA.
 XX AAW28912;
 AC
 XX 20-JAN-1998 (first entry)
 DT Opioid peptide.
 DE
 DE enkephalin; mu-opioid receptor ligand; agonist; antagonist.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 PH Modified-site 1 /note= "N-acetyl-Arg"
 FT Modified-site 6
 FT /note= "the C-terminal is in amide form"
 FT
 XX US5641861-A.
 XX 24-JUN-1997.
 PD 07-JUN-1995; 95US-00487006.
 PF 07-JUN-1995; 95US-00487006.
 PR
 XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.
 FA Houghten RA, Dooley CT;
 XX WPI; 1997-340994/31.
 DR
 XX New opioid peptide(s) which bind mu receptors specifically - have agonist
 PT or antagonist activity and are used for study and localisation of mu
 PT receptors and to treat peripheral side effects of morphine etc.

PS Disclosure; Col 8; 92pp; English.

XX The patent discloses the following new peptides, which are opioids which

XX bind specifically to the mu receptor: Ac-Phe-Arg-Trp-Trp-Tyr-Xaa-NH₂ (1);

CC Ac-Arg-Trp-Ile-Gly-Trp-Xaa-NH₂ (2); Trp-Trp-Pro-Lys-His-Xaa-NH₂ (3); Trp-

CC Trp-Pro-Xaa-NH₂ (4); Tyr-Pro-Phe-Gly-Phe-Xaa-NH₂ (5); D-Ile-D-Met-D-Ser-

CC D-Trp-D-Trp-(Gly)n-Xaa-NH₂ (6); D-Ile-D-Met-D-Trp-D-Trp-Gly-Xaa-NH₂ (7)

CC ; Tyr-Al-B₂-C₃-NH₂ (214); Pm and red (Me)x(H)y-Tyr-(NMe)z-Tyr-(Xaa3)z-

CC NH₂ (221); and Trp-Trp-Pro-D₄-(His)z-(Xaa)z-NH₂ (222); where Xaa = any

CC natural amino acid; Xaa1 = Lys or Arg; n and z = 0 or 1; Xaa2 = Gly or

CC the D form of any naturally occurring amino acid; Al = D-norvaline or D-

CC norleucine; B₂ = Gly, Phe or Trp; C₃ = Trp or naphthylalanine; x and y =

CC 0-2, but not over 2 in total; Xaa3 = Phe, Dphe or benzylamino; D₄ = Lys

CC or Arg; Pm and red indicate permethylation and reduction of all CO in

CC peptide links to methylene. These new compounds are useful: (i) for in

CC vitro assay and study of opiate receptor subtypes, particularly mu

CC receptors in the brain; (ii) for in vivo localisation of receptor

CC subtypes; and (iii) therapeutically to block the peripheral effects (e.g.

CC constipation and pruritus) of centrally acting pain killers such as

CC morphine. They are very selective for the mu opioid receptor, over

CC binding to the delta and kappa receptor subtypes. The present sequence is

CC a specific example of a peptide (2)

XX

SQ Sequence 6 AA;

Query Match 68.9%; Score 31; DB 2; Length 6;

Best Local Similarity 60.0%; Pred. No. 1.4e+06;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5

Db 2 WIGWH 6

RESULT 12

AAR93770

ID AAR93770 standard; protein; 6 AA.

AC AAR93770;

XX

XX 23-SEP-1997 (first entry)

DT

XX

XX New peptide which acts as mu-opioid receptor ligand.

DE

XX mu-receptor; opioid; opiate; agonist; antagonist; diagnosis; analgesic.

XX

XX Synthetic.

OS

XX

XX Key Location/Qualifiers

FH Modified-site 1

FT /note= "N-acetyl-Arg"

FT

FT Misc-difference 6

FT /note= "this residue is in C-terminal amide form"

XX

XX WO9640208-A1.

PN

XX

XX 19-DEC-1996.

PD

XX

XX 06-JUN-1996; 96WO-US009321.

PF

XX

XX 07-JUN-1995; 95US-00476438.

PR

XX

XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.

PA

XX

XX Dooley CT, Houghten RA;

PI

XX

XX WPI; 1997-051895/05.

DR

XX

XX New mu opioid receptor binding ligand peptide(s) - useful for in-vitro

PT and in-vivo diagnosis, as analgesics, and for blocking peripheral effects

PT of centrally acting drugs, e.g. morphine.

XX

PS Disclosure; Page 19; 57pp; English.

XX The patent discloses eight new groups of opioid peptides which bind to

CC the mu-receptor to act as agonists or antagonists. The peptides can be

CC used for in-vitro assays to study opiate receptor subtypes (especially

CC the mu type) in brain or other tissue samples; and for in-vivo diagnosis

CC to localise opioid subtypes. The peptides are also useful as drugs to

CC treat pathologies associated with other compounds which interact with the

CC opioid receptor system. Therefore they can be used in medicaments for

CC treating pathologies associated with the mu receptor and as analgesics.

CC They can be used therapeutically to block the peripheral effects of

CC centrally acting pain killers, e.g. to prevent side effects such as

CC constipation and pruritis associated with morphine. The present sequence

CC represents a specific example of one of the new groups of peptides, of

CC formula Ac-Arg-Trp-Ile-Gly-Trp-Xaa-NH₂ where Xaa = a naturally occurring

CC amino acid

XX

SQ Sequence 6 AA;

Query Match 68.9%; Score 31; DB 2; Length 6;

Best Local Similarity 60.0%; Pred. No. 1.4e+06;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5

Db 2 WIGWH 6

RESULT 13

AAV23019

ID AAV23019 standard; peptide; 6 AA.

XX

XX AAV23019;

AC

XX

XX 23-AUG-1999 (first entry)

DT

XX

XX Opioid peptide which inhibits binding of enkephalin.

DE

XX

XX Opioid peptide; ligand binding; opioid receptor;

KW micro-selective opioid peptide; enkephalin; opioid receptor system;

XX blocking; peripheral effect; centrally acting pain killer; morphine.

XX

XX Synthetic.

OS

XX

XX Key Location/Qualifiers

FH Modified-site 1

FT /note= "acetylated"

FT

FT Modified-site 6

FT /note= "amidated"

XX

XX US5919897-A.

PN

XX

XX 06-JUL-1999.

PD

XX

XX 07-JUN-1995; 95US-00488659.

PF

XX

XX 07-JUN-1995; 95US-00488659.

PR

XX

XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.

PA

XX

XX Dooley CT, Houghten RA;

PI

XX

XX WPI; 1999-394647/33.

DR

XX

XX New opioid peptides useful for blocking the peripheral effects of

PT centrally acting pain killers such as morphine.

PT

XX

XX Example 1; Col 8; 92pp; English.

PS

XX

XX The specification describes opioid peptides, in which each of the N atoms

CC in the peptide backbone between respective amino acids is modified by

CC permethylation, perallylation, perethylation, perbenzylation and

CC pernapththylation. The peptides inhibit ligand binding to an opioid

CC receptor. Specifically, the peptides inhibit the micro-selective opioid
 CC peptide enkephalin. The peptides can be used in vivo diagnostically to
 CC localize opioid receptor subtypes. They can be used to treat pathologies
 CC associated with other compounds which interact with the opioid receptor
 CC system. The peptides are especially useful for blocking the peripheral
 CC effects of centrally acting pain killers such as morphine. AAY23005-
 CC Y23024 represent opioid peptides of the invention, and are derived from
 CC the general sequence given in AAY23004
 XX
 XX Sequence 6 AA;

Query Match 68.9%; Score 31; DB 2; Length 6;
 Best Local Similarity 60.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5
 Db 2 WIGWH 6

RESULT 14
 AAB01507
 ID AAB01507 standard; peptide; 6 AA.
 XX
 AC AAB01507;
 XX
 DT 08-NOV-2000 (first entry)
 XX

Peptide which binds to transcription factor E2F-1 DNA binding domain.

XX DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
 KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
 KW restenosis.
 XX
 OS Synthetic.
 XX

PN WO200044771-A1.

XX 03-AUG-2000.

PF 26-JAN-2000; 2000WO-GB000227.

PR 26-JAN-1999; 99GB-00001710.

XX (PROL-) PROLIFIX LTD.

XX Mueller R, Kontermann RE, Montigiani S;

DR WPI; 2000-532806/48.

XX Peptides binding to the DNA binding domain of transcription factor E2F
 PT and inhibiting cell cycle progression, useful for the treatment of
 PT cancer.

XX Example; Page 26; 42pp; English.

XX Peptides which bind to the DNA binding domain of transcription factor E2F
 CC and inhibit cell cycle progression may be useful as research agents to
 CC investigate the interaction between E2F and DP-1, or the activation of
 CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
 CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
 CC treatment of cancer or other proliferative disorders such as psoriasis
 CC and restenosis
 XX

XX Sequence 6 AA;

Query Match 68.9%; Score 31; DB 3; Length 6;
 Best Local Similarity 83.3%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 Db 1 WVRWHF 6

RESULT 15
 AAM45777
 ID AAM45777 standard; peptide; 7 AA.
 XX
 AC AAM45777;
 XX

DT 25-OCT-2001 (first entry)
 XX

DE H11 binding site consensus conforming peptide (CCP) #2048.

XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
 KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
 KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
 KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
 KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
 XX

OS Homo sapiens.

OS Synthetic.

XX CA2290722-A1.

XX 08-JUN-2001.

PF 08-DEC-1999; 99CA-02290722.

PR 08-DEC-1999; 99CA-02290722.

XX (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

PI Entwistle JM, Macdonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress
 PT protein-peptide complexes associated with tumor, and isolated antigen-
 PT binding fragments of an antibody that binds specifically to the complex.

XX Example 4; Page 108; 154pp; English.

XX The present invention describes a composition (I) comprising stress
 CC protein-peptide complexes (SPPC) associated with tumours that is
 CC specifically immunogenically cross-reactive with cell surface-associated
 CC SPPCs specific to target cancer (TC). Also described is an isolated
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs
 CC or a population of different SPPCs consisting of immunogenic cancer cell
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
 CC used in vaccine production and as a tumour-specific immunogenic response
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
 CC or imaging cancer cells, and to monitor the course of amelioration of
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
 CC which are used in the exemplification of the present invention

XX Sequence 7 AA;

Query Match 68.9%; Score 31; DB 4; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWHF 6
 Db 1 WVRWHF 6

Search completed: June 10, 2004, 11:06:17
 Job time : 47.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:52:15 ; Search time 11.333 Seconds
(without alignments)
50.925 Million cell updates/sec

Title: US-09-912-414-2

Perfect score: 45

Sequence: 1 WVRWHF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 2522

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	60.0	10	F49033	T-cell receptor ga
2	24	53.3	10	T14019	cytochrome-c oxida
3	23	51.1	9	D57444	neuropeptide Grb-A
4	22	48.9	9	S07241	litorin - Rohde's
5	21	46.7	8	T13818	cytochrome oxidase
6	21	46.7	10	T13976	cytochrome-c oxida
7	21	46.7	10	T17057	cytochrome-c oxida
8	21	46.7	10	T12303	cytochrome-c oxida
9	21	46.7	10	T17060	cytochrome-c oxida
10	21	46.7	10	T14043	cytochrome-c oxida
11	21	46.7	10	T14054	cytochrome-c oxida
12	21	46.7	10	T17066	cytochrome-c oxida
13	21	46.7	10	T17069	cytochrome-c oxida
14	21	46.7	10	T12308	cytochrome-c oxida
15	21	46.7	10	T17072	cytochrome-c oxida
16	21	46.7	10	T12312	cytochrome-c oxida
17	21	46.7	10	T12316	cytochrome-c oxida
18	21	46.7	10	T12321	cytochrome-c oxida
19	21	46.7	10	T14219	cytochrome-c oxida
20	21	46.7	12	PH1308	Ig heavy chain DJ
21	20	44.4	9	S07205	litorin 2-Glu - Au
22	20	44.4	9	S07204	litorin I - Austr
23	20	44.4	9	A43848	cell surface adhes
24	20	44.4	11	C59151	protein-tyrosine k
25	20	44.4	12	PH1675	Ig heavy chain V r
26	20	44.4	13	PH1676	Ig heavy chain V r
27	20	44.4	13	A60409	bombesin-like pept
28	20	44.4	13	S23372	T-cell receptor al
29	20	44.4	14	PH1677	Ig heavy chain V r

30 20 44.4 14 2 PH1705
31 19 42.2 10 2 T17054
32 19 42.2 10 2 T17063
33 19 42.2 10 2 T12325
34 19 42.2 10 2 T12329
35 19 42.2 10 2 T14215
36 19 42.2 10 2 T14223
37 19 42.2 11 2 A49037
38 19 42.2 12 2 PH1324
39 18 40.0 10 2 B3887
40 17.5 38.9 10 2 T17075
41 17 37.8 5 2 A60803
42 17 37.8 11 2 S45698
43 17 37.8 11 2 A33571
44 17 37.8 12 2 A29169
45 17 37.8 12 2 A42324

ALIGNMENTS

RESULT 1

F49033

T-cell receptor gamma chain V-D-J region - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 19-Dec-1993 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000

C:Accession: F49033

R:Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.

Eur. J. Immunol. 21, 2999-3007, 1991

A:Title: Functionally distinct subsets of human gamma/delta T cells.

A:Reference number: A49033; MUID:92083926; PMID:1684157

A:Accession: F49033

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-10 <MOR>

A:Cross-references: GB:S72605; NID:g240700; PIDN:AA20632.1; PID:g240701

A>Note: sequence extracted from NCBI backbone (NCBIN:72605, NCBIP:72606)

C:Keywords: T-cell receptor

Query Match 60.0%; Score 27; DB 2; Length 10;

Best Local Similarity 60.0%; Pred. No. 1.1e+02;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5

DB |||:

4 WERWY 8

RESULT 2

T14019

cytochrome-c oxidase (EC 1.9.3.1) chain I - Eremias grammica mitochondrion (fragment)

C:Species: mitochondrion Eremias grammica

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000

C:Accession: T14019

R:Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.

Mol. Biol. Evol. 14, 91-104, 1997

A:Title: Two novel gene orders and the role of light-strand replication in rearrangeme

A:Reference number: Z17789; MUID:97153826; PMID:9000757

A:Accession: T14019

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-10 <MAC>

A:Cross-references: EMBL:U71331; NID:gl753240; PID:gl753243; PIDN:AA48277.1

C:Genetics:

A:Genome: mitochondrion

A>Note: COI

C:Keywords: mitochondrion; oxidoreductase

Query Match 53.3%; Score 24; DB 2; Length 10;

Best Local Similarity 60.0%; Pred. No. 3.2e+02;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VRWHF 6

Db 4 IRWFF 8

RESULT 3

D57444
neuropeptide Grb-AST B4 - two-spotted cricket
C:Species: Gryllus bimaculatus (two-spotted cricket)
C:Date: 26-Jan-1996 #sequence_revision 26-Jan-1996 #text_change 26-Jan-1996
C:Accession: D57444

R:Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.

J. Biol. Chem. 270, 21103-21108, 1995

A:Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cricket

A:Reference number: A57444; MUID:95403341; PMID:7673141

A:Accession: D57444

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <LOR>

Query Match 51.1%; Score 23; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WVRWH 5

Db 2 WERFH 6

RESULT 4

S07241
litorin - Rohde's leaf frog
C:Species: Phyllomedusa rohdei (Rohde's leaf frog)
C:Date: 12-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 18-Aug-2000
C:Accession: S07241

R:Barra, D.; Falconieri Erspamer, G.; Simmaco, M.; Bossa, F.; Melchiorri, P.; Erspamer,

PES Lett. 182, 53-56, 1985

A:Title: Rohdei-litorin: a new peptide from the skin of Phyllomedusa rohdei.

A:Reference number: S07241; MUID:85127560; PMID:3838283

A:Accession: S07241

A:Molecule type: protein

A:Residues: 1-9 <BAR>

C:Superfamily: gastrin-releasing peptide

C:Keywords: amidated carboxyl end; blocked amino end; neuropeptide; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:9/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 48.9%; Score 22; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WVRWHF 6

Db 3 WATGHF 8

RESULT 5

T13818
cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)
C:Species: Mitochondrion Myxine glutinosa (Atlantic hagfish)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13818

R:Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.

Mol. Biol. Evol. 14, 807-813, 1997

A:Title: The main features of the cranial mitochondrial DNA between the ND1 and the COI

A:Reference number: Z17775; MUID:97398704; PMID:9254918

A:Accession: T13818

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-8

A:Cross-references: EMBL:Y09527; NID:G2340019; PIDN:CAA70718.1; PID:G2340022

C:Genetics:

A:Genome: mitochondrion

A>Note: COI

C:Keywords: mitochondrion

Query Match 46.7%; Score 21; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWHF 6

Db 5 RWFF 8

RESULT 6

T13976
cytochrome-c oxidase (EC 1.9.3.1) chain I - Cnemidophorus tigris mitochondrion (fragment)
C:Species: mitochondrion Cnemidophorus tigris
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T13976

R:Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.

Mol. Biol. Evol. 14, 91-104, 1997

A:Title: Two novel gene orders and the role of light-strand replication in rearrangement

A:Reference number: Z17789; MUID:97153826; PMID:9000757

A:Accession: T13976

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-10 <MAC>

A:Cross-references: EMBL:U71332; NID:G1753236; PID:G1753239; PIDN:AAH48274.1

C:Genetics:

A:Genome: mitochondrion

A>Note: COI

C:Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 9.6e+02;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWHF 6

Db 5 RWFF 8

RESULT 7

T17057
cytochrome-c oxidase (EC 1.9.3.1) chain I - Crotaphytus collaris mitochondrion (fragment)
C:Species: mitochondrion Crotaphytus collaris
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 22-Oct-1999
C:Accession: T17057

R:Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.

J. Mol. Evol. 44, 660-674, 1997

A:Title: Evolutionary shifts in three major structural features of the mitochondrial

A:Reference number: Z18674; MUID:97315309; PMID:9169559

A:Accession: T17057

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-10 <MAC>

A:Cross-references: EMBL:U82681; NID:G3603108; PID:G3603111; PIDN:AAC62272.1

C:Genetics:

A:Genome: mitochondrion

A>Note: COI

C:Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 9.6e+02;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWHF 6

Db 5 RWFF 8

RESULT 8

T12303
cytochrome-c oxidase (EC 1.9.3.1) chain I - Diposaurus dorsalis mitochondrion (fragment)
C:Species: mitochondrion Diposaurus dorsalis

C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 22-Oct-1999

C;Accession: T12303
R;Schulte, J.A.; Macey, J.R.; Larson, A.; Papenfuss, T.J.
Mol. Phylogenet. Evol. 10, 367-376, 1998

A;Title: Molecular tests of phylogenetic taxonomies: A general procedure and example using

A;Reference number: Z17488; MUID:99162288; PMID:10051389

C;Accession: T12303

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-10 <SCH>

A;Cross-references: EMBL:AF049857; NID:g4105726; PID:g4105729; PIDN:AAD02514.1

C;Genetics:

A;Genome: mitochondrion

A;Note: COI

C;Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 9.6e+02;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWHF 6

DB 5 RWFF 8

RESULT 9

T17060

cytochrome-c oxidase (EC 1.9.3.1) chain I - Gambelia wislizenii mitochondrion (fragment)

C;Species: mitochondrion Gambelia wislizenii

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 22-Oct-1999

C;Accession: T17060

R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.

J. Mol. Evol. 44, 660-674, 1997

A;Title: Evolutionary shifts in three major structural features of the mitochondrial gene

A;Reference number: Z18674; MUID:97315309; PMID:9169559

C;Accession: T17060

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-10 <MAC>

A;Cross-references: EMBL:U82682; NID:g3603120; PID:g3603123; PIDN:AAC62281.1

C;Genetics:

A;Genome: mitochondrion

A;Note: COI

C;Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 9.6e+02;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWHF 6

DB 5 RWFF 8

RESULT 10

T14043

cytochrome-c oxidase (EC 1.9.3.1) chain I - Lialis jicari mitochondrion (fragment)

C;Species: mitochondrion Lialis jicari

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000

C;Accession: T14043

R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.

Mol. Biol. Evol. 14, 91-104, 1997

A;Title: Two novel gene orders and the role of light-strand replication in rearrangement

A;Reference number: Z17789; MUID:97153826; PMID:9000757

C;Accession: T14043

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-10 <MAC>

A;Cross-references: EMBL:U71327; NID:g1753244; PID:g1753247; PIDN:ARB48280.1

C;Genetics:

A;Genome: mitochondrion

A;Note: COI

C;Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 9.6e+02;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWHF 6

DB 5 RWFF 8

RESULT 11

T14054

cytochrome-c oxidase (EC 1.9.3.1) chain I - Mabuya aurata mitochondrion (fragment)

C;Species: mitochondrion Mabuya aurata

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000

C;Accession: T14054

R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.

Mol. Biol. Evol. 14, 91-104, 1997

A;Title: Two novel gene orders and the role of light-strand replication in rearrangement

A;Reference number: Z17789; MUID:97153826; PMID:9000757

C;Accession: T14054

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-10 <MAC>

A;Cross-references: EMBL:U71330; NID:g1753248; PID:g1753251; PIDN:ARB48283.1

C;Genetics:

A;Genome: mitochondrion

A;Note: COI

C;Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 9.6e+02;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWHF 6

DB 5 RWFF 8

RESULT 12

T17066

cytochrome-c oxidase (EC 1.9.3.1) chain I - Oplurus cuvieri mitochondrion (fragment)

C;Species: mitochondrion Oplurus cuvieri

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 22-Oct-1999

C;Accession: T17066

R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.

J. Mol. Evol. 44, 660-674, 1997

A;Title: Evolutionary shifts in three major structural features of the mitochondrial ge

A;Reference number: Z18674; MUID:97315309; PMID:9169559

C;Accession: T17066

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-10 <MAC>

A;Cross-references: EMBL:U82685; NID:g3603136; PID:g3603139; PIDN:AAC62293.1

C;Genetics:

A;Genome: mitochondrion

A;Note: COI

C;Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 9.6e+02;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWHF 6

DB 5 RWFF 8

RESULT 13

T17069

cytochrome-c oxidase (EC 1.9.3.1) chain I - Phrynosoma douglassii mitochondrion (fragme

C;Species: mitochondrion Phrynosoma douglassii

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 22-Oct-1999

C;Accession: T17069
 R;Macey, J.R.; Larson, A.; Ananijeva, N.B.; Papenfuss, T.J.
 J. Mol. Evol. 44, 660-674, 1997
 A;Title: Evolutionary shifts in three major structural features of the mitochondrial gene
 A;Reference number: Z18674; MUID:97315309; PMID:9169559
 A;Accession: T17069
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: DNA
 A;Residues: 1-10 <MAC>
 A;Cross-references: EMBL:U82686; NID:G3603144; PID:G3603147; PIDN:AAC62299.1
 C;Genetics:
 A;Genome: mitochondrion
 A;Note: COI
 C;Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;
 Best Local Similarity 75.0%; Pred. No. 9.6e+02;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 RWHF 6
 Db 5 RWFF 8

RESULT 14

T12308
 cytochrome-c oxidase (EC 1.9.3.1) chain I - Sator angustus mitochondrion (fragment)
 C;Species: mitochondrion Sator angustus
 C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 22-Oct-1999
 C;Accession: T12308
 R;Schulte, J.A.; Macey, J.R.; Larson, A.; Papenfuss, T.J.
 Mol. Phylogenet. Evol. 10, 367-376, 1998
 A;Title: Molecular tests of phylogenetic taxonomies: A general procedure and example usi
 A;Reference number: Z17488; MUID:99162288; PMID:10051389
 A;Accession: T12308
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: DNA
 A;Residues: 1-10 <SCH>
 A;Cross-references: EMBL:AF049859; NID:G4105734; PID:G4105737; PIDN:AAD02520.1
 C;Genetics:
 A;Genome: mitochondrion
 A;Note: COI
 C;Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;
 Best Local Similarity 75.0%; Pred. No. 9.6e+02;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 RWHF 6
 Db 5 RWFF 8

RESULT 15

T17072
 cytochrome-c oxidase (EC 1.9.3.1) chain I - Sauromalus obesus mitochondrion (fragment)
 C;Species: mitochondrion Sauromalus obesus
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 22-Oct-1999
 C;Accession: T17072
 R;Macey, J.R.; Larson, A.; Ananijeva, N.B.; Papenfuss, T.J.
 J. Mol. Evol. 44, 660-674, 1997
 A;Title: Evolutionary shifts in three major structural features of the mitochondrial ge
 A;Reference number: Z18674; MUID:97315309; PMID:9169559
 A;Accession: T17072
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: DNA
 A;Residues: 1-10 <MAC>
 A;Cross-references: EMBL:U82687; NID:G3603152; PID:G3603155; PIDN:AAC62305.1
 C;Genetics:
 A;Genome: mitochondrion
 A;Note: COI
 C;Keywords: mitochondrion; oxidoreductase

Query Match 46.7%; Score 21; DB 2; Length 10;
 Best Local Similarity 75.0%; Pred. No. 9.6e+02;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 RWHF 6
 Db 5 RWFF 8

Search completed: June 10, 2004, 11:07:04
 Job time : 12.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 11:01:15 ; Search time 7.66667 Seconds
(without alignments)
40.751 Million cell updates/sec

Title: US-09-912-414-2

Perfect score: 45

Sequence: 1 WVRWHF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 801

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	48.9	9	1 LITR_PHYRO	P08946 phyllomedusa
2	22	48.9	11	1 RANC_RANPI	P08951 rana pipien
3	20	44.4	9	1 LITO_LITAU	P08945 litoria aur
4	20	44.4	13	1 BOML_PSEGU	P42991 pseudophryn
5	19.5	43.3	5	1 UF01_MOUSE	P38639 mus musculus
6	19	42.2	10	1 LABA_JATMU	P3270 jatropha mu
7	18	40.0	9	1 COW_CONVE	P83047 conus ventr
8	18	40.0	13	1 YENP_PHOLU	P41122 photorhabdu
9	17	37.8	11	1 MLG_THERS	P41989 theromyzon
10	17	37.8	13	1 EI21_LITRU	P82097 litoria rub
11	17	37.8	13	1 EI22_LITRU	P82098 litoria rub
12	17	37.8	13	1 TEML_RANTE	P57104 rana tempor
13	16	35.6	8	1 ACI_THUAL	P18691 thunnus alb
14	16	35.6	13	1 MIA_ANOCA	P41589 anolis caro
15	16	35.6	13	1 MIA_CAMOR	P01198 camelus dro
16	16	35.6	14	1 LPW_RHIME	P18854 rhizobium m
17	15	33.3	7	1 TPFY_PACDA	P83455 pachymedusa
18	15	33.3	10	1 AEGL_AGRAE	P83465 agroclybe ae
19	15	33.3	12	1 RF1_CONSP	P58805 conus spuri
20	15	33.3	15	1 CX3B_CONQU	P58842 conus querc
21	15	33.3	15	1 GLN2_PINPS	P81107 pinus pinas
22	14	31.1	10	1 BPP2_BOTJA	P01022 bothrops ja
23	14	31.1	10	1 FARP_MYTED	P42560 mytilus edu
24	14	31.1	10	1 GRP_RANPI	P23260 rana ridibu
25	14	31.1	11	1 CA22_LITCI	P82088 litoria cit
26	14	31.1	11	1 CA42_LITCI	P82092 litoria cit
27	14	31.1	11	1 CA1F_SARBU	P83518 sarcophaga
28	14	31.1	13	1 BPPI_BOTJA	P01020 bothrops ja
29	14	31.1	13	1 CXA2_CONGE	P01520 conus geogr
30	14	31.1	13	1 ALYT_ALYOB	P08944 alytes obst
31	14	31.1	14	1 MAST_PARID	P42716 parapolybia
32	14	31.1	14	1 MAST_VESPA	P21654 vespa basal
33	14	31.1	14	1 MAST_VESXA	P01515 vespa xanth

34 14 31.1 15 1 AH2_PRUSE
35 14 31.1 15 1 DCMW_PSECH
36 14 31.1 15 1 MK2A_PALPR
37 14 31.1 15 1 RM12_YEAST
38 13 28.9 5 1 BPP7_BOTIN
39 13 28.9 9 1 NEF_HV1Z8
40 13 28.9 10 1 APE_CAPGI
41 13 28.9 10 1 GON1_ALIMI
42 13 28.9 10 1 GON2_CHEPR
43 13 28.9 10 1 GON2_CHICK
44 13 28.9 10 1 GON3_ONCKE
45 13 28.9 10 1 NO40_TOBAC

P29260 prunus sero
P19917 pseudomonas
P80409 palomena pr
P36522 saccharomyc
P30425 bothrops in
P12481 human immu
P80474 capnocytoph
P37041 alligator m
P80678 chelyosoma
P37043 gallus gall
P20367 oncorhynchu
P55962 nicotiana t

ALIGNMENTS

RESULT 1
LITR_PHYRO STANDARD; PRT; 9 AA.
AC P08946;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rhodai-litorin.
OS Phyllomedusa rohdei (Rohde's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;
OC Phyllomedusinae; Phyllomedusa.
OX NCBI_TaxID=8394;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=85127560; PubMed=3838283;
RA Barra D., Espamer G.F., Simmaco M., Bossa F., Melchiorri P.,
RA Espamer V.;
RT "Rohdei-litorin: a new peptide from the skin of Phyllomedusa rohdei.";
RL FEBS Lett. 182:53-56(1985)
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Skin.
CC -1- SIMILARITY: Belongs to the bombesin/neuromedin B/ratanensin family.
CC PIR; S07241; S07241.
DR InterPro; IPR000874; Bombesin.
DR Pfam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
KW Amphibian defense peptide; Bombesin family; Amidation;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1090 MW; 4ECCC18861ADC377 CRC64;

Query Match 48.9%; Score 22; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 3 WATGHF 8

RESULT 2

RANC_RANPI STANDARD; PRT; 11 AA.
ID RANC_RANPI
AC P08951;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ranatensin-C.
OS Rana pipiens (Northern leopard frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Ranidae; Rana.
OX NCBI_TaxID=8404;

[1]
 RN SEQUENCE.
 RP TISSUE=Skin secretion;
 RC MEDLINE=84131098; PubMed=6141890;
 RA Nakajima T.;
 RL Unpublished results, cited by:
 RL Erspamer V., Erspamer G.F., Mazzanti G., Eudean R.;
 RL Comp. Biochem. Physiol. 77C:99-108(1984).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Skin.
 CC -1- SIMILARITY: Belongs to the bombesin/neuromedin B/ranatensin family.
 DR InterPro; IPR000874; Bombesin.
 DR Pfam; PF02044; Bombesin; 1.
 DR PROSITE; PS00257; BOMBESIN; 1.
 KW Amphibian defense peptide; Bombesin family; Amidation.
 FT MOD_RES 11 11
 SQ SEQUENCE 11 AA; 1304 MW; D6C9885A61ADC366 CRC64;
 Query Match 48.9%; Score 22; DB 1; Length 11;
 Best Local Similarity 50.0%; Pred. No. 2e+02; Indels 3; Gaps 0;
 Matches 3; Conservative 0; Mismatches 3; Indels 3; Gaps 0;
 QY 1 WVRWHF 6
 DB 5 WATGHF 10

RESULT 3
 LITO_LITAU STANDARD; PRT; 9 AA.
 AC P08945;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Litorin.
 OS Litoria aurea (Green and golden bell frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaidea; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=8371;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RX MEDLINE=75187011; PubMed=1140241;
 RA Anastasi A., Erspamer V., Eudean R.;
 RT "Aminoacid composition and sequence of litorin, a bombesin-like nonapeptide from the skin of the Australian leptodactylid frog Litoria aurea";
 RL Litoria aurea";
 RL Experientia 31:510-511(1975).
 RN [2]
 RP SEQUENCE, AND METHYLATION OF GLN-2.
 RC TISSUE=Skin secretion;
 RX MEDLINE=78003546; PubMed=908397;
 RA Anastasi A., Montecucchi P.C., Angelucci F., Erspamer V., Eudean R.;
 RT "Glu(OMe)3-litorin, the second bombesin-like peptide occurring in methanol extracts of the skin of the Australian frog Litoria aurea";
 RL Experientia 33:1289-1289(1977).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Skin.
 CC -1- SIMILARITY: Belongs to the bombesin/neuromedin B/ranatensin family.
 DR PIR; S07204; S07204.
 DR PIR; S07205; S07205.
 DR InterPro; IPR000874; Bombesin.
 DR Pfam; PF02044; Bombesin; 1.
 DR PROSITE; PS00257; BOMBESIN; 1.
 KW Amphibian defense peptide; Bombesin family; Amidation; Methylation;
 KW Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1
 FT MOD_RES 2 2
 FT MOD_RES 9 9
 SQ SEQUENCE 9 AA; 1103 MW; D7CCC1E862CDC366 CRC64;

Query Match 44.4%; Score 20; DB 1; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1.4e+05;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WVRWHF 6
 DB 3 WATGHF 8

RESULT 4
 BOML_PSEGU STANDARD; PRT; 13 AA.
 AC P42991;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bombesin-like peptide 1 (pg-L).
 OS Pseudophryne guentheri (Guenther's toadlet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaidea; Myobatrachidae;
 OC Myobatrachinae; Pseudophryne.
 OX NCBI_TaxID=30349;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RX MEDLINE=90287814; PubMed=2356157;
 RA Simmaco M., Severini C., de Biase D., Barra D., Bossa F., Roberts J.D., Melchiorri P., Erspamer V.;
 RT "Six novel tachykinin- and bombesin-related peptides from the skin of the Australian frog Pseudophryne guentheri";
 RL Peptides 11:299-304(1990).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Skin.
 CC -1- SIMILARITY: Belongs to the bombesin/neuromedin B/ranatensin family.
 DR PIR; A60409; A60409.
 DR InterPro; IPR000874; Bombesin.
 DR Pfam; PF02044; Bombesin; 1.
 DR PROSITE; PS00257; BOMBESIN; 1.
 KW Amphibian defense peptide; Bombesin family; Amidation;
 KW Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1
 FT MOD_RES 13 13
 SQ SEQUENCE 13 AA; 1372 MW; D6DE0D24BD98C366 CRC64;
 Query Match 44.4%; Score 20; DB 1; Length 13;
 Best Local Similarity 50.0%; Pred. No. 5e+02; Indels 3; Gaps 0;
 Matches 3; Conservative 0; Mismatches 3; Indels 3; Gaps 0;
 QY 1 WVRWHF 6
 DB 7 WATGHF 12

RESULT 5
 UF01_MOUSE STANDARD; PRT; 5 AA.
 AC P38639;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins

RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
 FT protein is: 6.6, its MW is: 19 kDa.
 CC NON_TER 5 5
 SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;
 Query Match 43.3%; Score 19.5; DB 1; Length 5;
 Best Local Similarity 60.0%; Pred. No. 1.4e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 WVRW 4
 Db | : ||
 1 WIGRW 5
 RESULT 6
 LABA JATMU STANDARD; PRT; 10 AA.
 ID LABA JATMU
 AC FL3270;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Labaditin.
 OS Jatropha multifida (Physic nut).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae; Jatrophaeae;
 CC Jatropha.
 CC NCBI_TaxID=3996;
 RN [1] _TaxID=3996;
 RP SEQUENCE.
 RC TISSUE=Lateax;
 RA Kosasi S., van der Sluis W.G., Boelens R., T'Hart L.A., Labadie R.P.;
 RT "Labaditin, a novel cyclic decapeptide from the latex of Jatropha
 RT multifida L. (Euphorbiaceae). Isolation and sequence determination
 RT by means of two-dimensional NMR.";
 RL FEBS Lett. 256:91-96(1989).
 CC -!- FUNCTION: Labaditin is an active peptide which inhibits the
 CC classical pathway of complement activation in vitro. Activity
 CC seems to be based on an interaction with C1.
 CC -!- PTM: This is a cyclic peptide.
 CC -!- MISCELLANEOUS: Latex of this plant is used in folkloric medicine
 CC for treatment of infected wounds, skin infections and scabies.
 SQ SEQUENCE 10 AA; 1089 MW; D98AAD6362D1B362 CRC64;
 Query Match 42.2%; Score 19; DB 1; Length 10;
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WVRW 4
 Db | : ||
 4 WTVW 7
 RESULT 7
 COW_CONV STANDARD; PRT; 9 AA.
 ID COW_CONV
 AC P83047;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Contryphan-Vn.
 OS Conus ventricosus (Mediterranean cone).
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 CC Apogastropoda; Caenogastropoda; Sorbeconcha; Hypogastropoda;
 CC Neogastropoda; Conoidea; Conidae; Conus.
 CC NCBI_TaxID=117992;
 RN [1]
 RP SEQUENCE, SYNTHESIS, DISULFIDE BONDS, AND MASS SPECTROMETRY.
 RC TISSUE=Venom;
 RX MEDLINE=21547785; PubMed=11688995;
 RA Massilia G.R., Schinina M.E., Ascenzi P., Politicelli F.;

RT "Contryphan-Vn: a novel peptide from the venom of the Mediterranean
 RT small Conus ventricosus.";
 RL Biochem. Biophys. Res. Commun. 288:908-913(2001).
 RN [2]
 RP STRUCTURE BY NMR, SYNTHESIS, DISULFIDE BONDS, AND FUNCTION.
 RX MEDLINE=22533239; PubMed=12646193;
 RA Massilia G.R., Eliseo T., Grolleau F., Lapiet B., Barbier J.,
 RA Bournaud R., Molgo J., Cicero D.O., Paci M., Schinina M.E.,
 RA Ascenzi P., Politicelli F.;
 RT "Contryphan-Vn: a modulator of Ca2+-dependent K+ channels.";
 RL Biochem. Biophys. Res. Commun. 303:238-246(2003).
 CC -!- FUNCTION: Affects both voltage-gated and calcium-dependent
 CC potassium channel activities, with composite and diversified
 CC effects in invertebrate and vertebrate systems.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- PTM: The cis isomer is the most abundant and is thus thought to be
 CC the functionally relevant conformer.
 CC -!- MASS SPECTROMETRY: MW=1088.6; METHOD=MALDI.
 CC -!- SIMILARITY: Belongs to the contryphan family.
 CC PDB; INXN; 04-MAR-03.
 KW Toxin; Ionic channel inhibitor; Neurotoxin;
 KW Potassium channel inhibitor; D-amino acid; Amidation; 3D-structure.
 FT DISULFID 3 9 D-TRYPTOPHAN.
 FT MOD_RES 5 5
 FT MOD_RES 9 9 AMIDATION.
 SQ SEQUENCE 9 AA; 1091 MW; 8D38676323676EBA CRC64;
 Query Match 40.0%; Score 18; DB 1; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1.4e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WVRW 4
 Db | : ||
 5 WKPW 8
 RESULT 8
 YPNP_PHOLU STANDARD; PRT; 13 AA.
 ID YPNP_PHOLU
 AC P41122;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein in pmp 3' region (ORF3) (Fragment).
 OS Photobacterium luminescens (Xenorhabdus luminescens).
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Photobacterium.
 CC NCBI_TaxID=29488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K122;
 RX MEDLINE=94266731; PubMed=8206856;
 RA Clarke D.J., Dowds B.C.A.;
 RT "The gene coding for polynucleotide phosphorylase in Photobacterium sp.
 RT strain K122 is induced at low temperatures.";
 RL J. Bacteriol. 176:3775-3784(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X76069; CAA53672.1; -
 KW Hypothetical protein.
 FT NON_TER 13
 SQ SEQUENCE 13 AA; 1634 MW; 64774A4F6267A364 CRC64;
 Query Match 40.0%; Score 18; DB 1; Length 13;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;

Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRW 4
Db 4 FLRW 7

RESULT 9
ID MLG THETS STANDARD; PRT; 11 AA.
AC P41389;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin gamma (Gamma-melanocyte stimulating hormone) (Gamma-MSH).
OS Theromyzon tessulatum (Leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Rhynchobdellida; Glossiphoniidae; Theromyzon.
OX NCBI_TaxID=13286;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RA MEDLINE=94298944; PubMed=8026574;
RX Salzet M., Watter C., Bulet P., Malecha J.;
RT "Isolation and structural characterization of a novel peptide related
RT to gamma-melanocyte stimulating hormone from the brain of the leech
RT Theromyzon tessulatum.";
RL FEBS Lett. 348:102-106(1994).
CC -!- SIMILARITY: Belongs to the POMC family.
KW PIR; S45698; S45698.
DR Hormone; Amidation.
FT MOD RES 11 11
SQ SEQUENCE 11 AA; 1486 MW; 2DB8FACE6409C1E8 CRC64;

Query Match 37.8%; Score 17; DB 1; Length 11;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWHF 6
Db 1 YVMGHF 6

RESULT 10
ID MLG THETS STANDARD; PRT; 13 AA.
AC P82097;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 2.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria rubella. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD RES 13 13
SQ SEQUENCE 13 AA; 1598 MW; C1808EF326F57322 CRC64;

Query Match 37.8%; Score 17; DB 1; Length 13;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRW 4
Db 6 VKW 8

RESULT 11
ID MLG THETS STANDARD; PRT; 13 AA.
AC P82098;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 2.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria rubella. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD RES 13 13
SQ SEQUENCE 13 AA; 1598 MW; C1808EF3B3B57322 CRC64;

Query Match 37.8%; Score 17; DB 1; Length 13;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRW 4
Db 6 VKW 8

RESULT 12
ID MLG THETS STANDARD; PRT; 13 AA.
AC P57104;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Temporin L.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=97175050; PubMed=9022710;
RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA Barra D.;
RT "Temporins, antimicrobial peptides from the European red frog Rana
RT temporaria.";
RL Eur. J. Biochem. 242:788-792(1996).
CC -!- FUNCTION: Has antibacterial activity against Gram-negative and
CC Gram-positive bacteria.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Belongs to the brevinin family.
KW Amphibian defense peptide; Antibiotic; Amidation.
FT MOD RES 13 13
SQ SEQUENCE 13 AA; 1641 MW; 9EBDCB1FAFF7C325 CRC64;

Query Match 37.8%; Score 17; DB 1; Length 13;

Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRW 4
Db : : :
1 FVQW 4

RESULT 13

ACI_THUAL STANDARD; PRT; 8 AA.
ID ACI_THUAL
AC P15691;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Angiotensin-converting enzyme inhibitor
OS Thunus albacares (Yellowfin tuna) (Neothunnus macropterus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8236;
RN [1]
RP SEQUENCE.
RC TISSUE=Muscle;
RX MEDLINE=8326322; PubMed=3415688;
RA Kohama Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;
RT "Isolation of angiotensin-converting enzyme inhibitor from tuna
muscle."
RL Biochem. Biophys. Res. Commun. 155:332-337 (1988).
DR PIR: A31570; A31570.
SQ SEQUENCE 8 AA; 953 MW; 6AA863733051F1B7 CRC64;

Query Match 35.6%; Score 16; DB 1; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.4e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRW 4
Db : : :
4 IKW 6

RESULT 14

MLA_ANOCA STANDARD; PRT; 13 AA.
ID MLA_ANOCA
AC P41589;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin alpha (Alpha-MSH).
OS Anolis carolinensis (Green anole) (American chameleon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Iguanidae; Polychrotinae; Anolis.
OX NCBI_TaxID=28377;
RN [1]
RP SEQUENCE.
RC TISSUE=Pituitary;
RX MEDLINE=92270473; PubMed=1667689;
RA Dore R.M., Lancha A., Rand-Weaver M., Jankelow L., Adamczyk D.L.;
RT "Detection of a novel sequence change in the major form of alpha-MSH
isolated from the intermediate pituitary of the reptile, Anolis
carolinensis."
RL Peptides 12:1261-1266 (1991).
CC -!- SIMILARITY: Belongs to the POMC family.
DR InterPro: IPR001941; Mcortin ACTH.
DR Pfam: PF00976; ACTH_domain; 1.
KW Hormone; Amidation.
FT MOD_RES 13 13
SQ SEQUENCE 13 AA; 1608 MW; FF990A7358BB09C1 CRC64;

Query Match 35.6%; Score 16; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RW 4
Db : : :
8 RW 9

RESULT 15

MLA_CAMDR STANDARD; PRT; 13 AA.
ID MLA_CAMDR
AC P01198;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin alpha (Alpha-MSH).
OS Camelus dromedarius (Dromedary) (Arabic camel), and
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838, 9796;
RN [1]
RP SEQUENCE.
RC SPECIES=Camelus dromedarius;
RX MEDLINE=75146434; PubMed=1125179;
RA Li C.H., Danho W.O., Chung D., Rao A.J.;
RT "Isolation, characterization, and amino acid sequence of
melanotropins from camel pituitary glands."
RL Biochemistry 14:947-952 (1975).
RN [2]
RP SEQUENCE.
RC SPECIES=Horse; TISSUE=Pituitary;
RA Dixon J.S., Li C.H.;
RT "The isolation and structure of alpha-melanocyte-stimulating hormone
from horse pituitaries."
RL J. Am. Chem. Soc. 82:4568-4572 (1960).
CC -!- SIMILARITY: Belongs to the POMC family.
DR PIR: A01464; MTCMD.
DR PIR: A91785; MTHOAR.
DR InterPro: IPR001941; Mcortin ACTH.
DR Pfam: PF00976; ACTH_domain; 1.
KW Hormone; Acetylation; Amidation.
FT MOD_RES 1 1
SQ SEQUENCE 13 AA; 1624 MW; FF991CA958BB09C1 CRC64;

Query Match 35.6%; Score 16; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RW 4
Db : : :
8 RW 9

Search completed: June 10, 2004, 11:07:41
Job time : 9.66667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 11:06:26 ; Search time 31.3333 Seconds
(without alignments)
60.418 Million cell updates/sec

Title: US-09-912-414-2

Perfect score: 45

Sequence: 1 WVRWHF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 3954

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTEMBL 25: **
2: sp_archaea: **
3: sp_bacteria: **
4: sp_fungi: **
5: sp_invertebrate: **
6: sp_mammal: **
7: sp_nhc: **
8: sp_organelle: **
9: sp_phase: **
10: sp_plant: **
11: sp_rodent: **
12: sp_virus: **
13: sp_vertebrate: **
14: sp_unclassified: **
15: sp_rvirus: **
16: sp_bacteriaph: **
17: sp_archaea: **

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	62.2	9	8 Q8SHF0	Q8shf0 chameleo n
2	24	53.3	8	8 Q94VF6	Q94vf6 varanus job
3	24	53.3	10	8 P92632	P92632 eremias gra
4	23	51.1	10	8 Q9TG41	Q9tg41 ophisaurus
5	22	48.9	8	8 Q94VJ4	Q94vj4 varanus ben
6	22	48.9	8	13 P79940	P79940 xenopus lae
7	22	48.9	10	8 Q9B4X0	Q9b4x0 notophthalm
8	22	48.9	10	8 Q958L2	Q958l2 rana tempor
9	22	48.9	10	8 Q958L8	Q958l8 rana catesb
10	22	48.9	10	8 Q958K6	Q958k6 rana pretio
11	22	48.9	10	8 Q958K0	Q958k0 rana casc
12	22	48.9	10	8 Q958L5	Q958l5 rana sylvat
13	22	48.9	10	8 Q958K3	Q958k3 rana aurora
14	22	48.9	10	8 Q94NH4	Q94nh4 rana muscos
15	22	48.9	10	8 Q94VD2	Q94vd2 varanus pan
16	21	46.7	8	8 Q94VC1	Q94vc1 varanus rud

17	21	46.7	8	8 Q9TD02	Q9td02 terranatos
18	21	46.7	8	8 Q9T4Y2	Q9t4y2 asterina pe
19	21	46.7	9	8 Q9T688	Q9t688 gecko gecko
20	21	46.7	10	8 Q9T8K7	Q9t8k7 liolaemus m
21	21	46.7	10	8 Q9T8N1	Q9t8n1 liolaemus p
22	21	46.7	10	8 Q9T903	Q9t903 oplurus cuv
23	21	46.7	10	8 Q8W969	Q8w969 anolis orto
24	21	46.7	10	8 Q8WDH8	Q8wdh8 anolis mees
25	21	46.7	10	8 Q9T8T6	Q9t8t6 liolaemus m
26	21	46.7	10	8 Q9T8L3	Q9t8l3 liolaemus l
27	21	46.7	10	8 P92616	P92616 aspidosceli
28	21	46.7	10	8 Q9T8G8	Q9t8g8 liolaemus c
29	21	46.7	10	8 Q958K9	Q958k9 rana boylii
30	21	46.7	10	8 Q9TFU9	Q9tfu9 teratoscinc
31	21	46.7	10	8 Q9T8X7	Q9t8x7 phymaturus
32	21	46.7	10	8 Q79885	Q79885 anolis pate
33	21	46.7	10	8 Q9T8O5	Q9t8o5 liolaemus l
34	21	46.7	10	8 P92654	P92654 euprepis au
35	21	46.7	10	8 Q9T8L0	Q9t8l0 liolaemus o
36	21	46.7	10	8 Q9T8W8	Q9t8w8 liolaemus b
37	21	46.7	10	8 Q9T8R4	Q9t8r4 liolaemus p
38	21	46.7	10	8 Q9T8M8	Q9t8m8 liolaemus m
39	21	46.7	10	8 Q9T8S1	Q9t8s1 liolaemus l
40	21	46.7	10	8 Q9T8S4	Q9t8s4 liolaemus c
41	21	46.7	10	8 Q9ZT04	Q9zt04 sceloporusc
42	21	46.7	10	8 P92758	P92758 teratoscinc
43	21	46.7	10	8 Q9T8T9	Q9t8t9 liolaemus l
44	21	46.7	10	8 Q9ZVT5	Q9zvt5 uta stansbu
45	21	46.7	10	8 Q9T8J8	Q9t8j8 liolaemus w

ALIGNMENTS

RESULT 1

Q8SHF0 ID Q8SHF0 PRELIMINARY; PRT; 9 AA.
AC Q8SHF0; 9 AA.
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Chamaeleo namaquensis (Namaqua chameleon).
OG Mitochondrion.
OC Lepidodactylus; Squamata; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Iguania; Acrodonta; Chamaeleonidae; Chamaeleo.
OX NCBI_TaxID=179917;
RN [1]
RP TOWNSEND T.M., LARSON A.L.;
RA "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
RT Chamaeleonidae (Reptilia, Squamata)."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448757; AAL90553.1;
DR GO; GO:0005739; C-mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1205 MW; 358CB72733640733 CRC64;

Query Match 62.2%; Score 28; DB 8; Length 9;
Best Local Similarity 75.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVRW 4
Db 2 WLRW 5

RESULT 2

Q94VF6 ID Q94VF6 PRELIMINARY; PRT; 8 AA.
AC Q94VF6;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Varanus jobiensis (Peach throat monitor).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
 OX NCBI_TaxID=169843;
 [1]
 RN SEQUENCE FROM N.A.
 RA Ast J.C.;
 RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
 RL Cladistics 17:0-0(2001).
 DR EMBL; AF407507; AAL10075.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER
 SQ SEQUENCE 8 AA; 1144 MW; EFD729DB436411A6 CRC64;
 Query Match 53.3%; Score 24; DB 8; Length 8;
 Best Local Similarity 75.0%; Pred. No. 1e+06;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RWHF 6
 Db ||| |
 3 RWYF 6

RESULT 3
 P92632
 ID P92632 PRELIMINARY; PRT; 10 AA.
 AC P92632;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Bremias grammica.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Lacertoidea;
 OC Lacertidae; Bremias.
 OX NCBI_TaxID=52179;
 [1]
 RN SEQUENCE FROM N.A.
 RA Macey J.R., Larson A., Ananjeva N.B., Fang Z., Papenfuss T.J.;
 RT "Two novel gene orders and the role of light-strand replication in
 rearrangement of the vertebrate mitochondrial genome.";
 RL Mol. Biol. Evol. 14:91-104(1997).
 [2]
 RN SEQUENCE FROM N.A.
 RA Macey J.R., Larson A., Ananjeva N.B., Papenfuss T.J.;
 RT "Replication slippage may cause parallel evolution in the secondary
 structures of mitochondrial transfer RNAs.";
 RL Mol. Biol. Evol. 14:30-39(1997).
 DR EMBL; U71331; AAB48277.1; -.
 DR PIR; T14019; T14019.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER
 SQ SEQUENCE 10 AA; 1288 MW; 5B3580C9D3640057 CRC64;
 Query Match 53.3%; Score 24; DB 8; Length 10;
 Best Local Similarity 60.0%; Pred. No. 8.4e+02;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VRWHF 6
 Db ||| |
 4 IRWFF 8

RESULT 4
 Q9TG41
 ID Q9TG41 PRELIMINARY; PRT; 10 AA.
 AC Q9TG41;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Ophisaurus apodus (Legless lizard).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguillidae;
 OC Ophisaurus.
 OX NCBI_TaxID=102191;
 [1]
 RN SEQUENCE FROM N.A.
 RA Macey J.R., Schulte J.A. II, Larson A., Tuniyev B.S., Orlov N.,
 RA Papenfuss T.J.;
 RT "Molecular phylogenetics, tRNA evolution, and historical biogeography
 in anguillid lizards and related taxonomic families.";
 RL Mol. Phylogenet. Evol. 12:250-272(1999).
 DR EMBL; AF085623; AAD51559.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER
 SQ SEQUENCE 10 AA; 1239 MW; 1A3580C7336412C0 CRC64;
 Query Match 51.1%; Score 23; DB 8; Length 10;
 Best Local Similarity 80.0%; Pred. No. 1.2e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VRWHF 6
 Db ||| |
 4 VRWLF 8

RESULT 5
 Q94VJ4
 ID Q94VJ4 PRELIMINARY; PRT; 8 AA.
 AC Q94VJ4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Varanus bengalensis nebulosis (Clouded monitor).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
 OX NCBI_TaxID=169827;
 [1]
 RN SEQUENCE FROM N.A.
 RA Ast J.C.;
 RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
 RL Cladistics 17:0-0(2001).
 DR EMBL; AF407492; AAL10031.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER
 SQ SEQUENCE 8 AA; 1053 MW; E8B5B9C733640056 CRC64;
 Query Match 48.9%; Score 22; DB 8; Length 8;
 Best Local Similarity 60.0%; Pred. No. 1e+06;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VRWHF 6
 Db ||| |
 2 IRWLF 6

```

RESULT 6
P79940
ID P79940 PRELIMINARY; PRT; 8 AA.
AC P79940;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE XMeis1-4 protein (Fragment).
OS Xenopus laevis (African clawed frog)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97202105; PubMed=9049632;
RA Steelman S., Moskow J.U., Muzynski K., North C., Druck T.,
RA Montgomery J.C., Huebner K., Daar I.O., Buchberg A.M.;
RT "Identification of a conserved family of Meis1-related homeobox
genes.";
RL Genome Res. 7:142-156(1997).
DR EMBL: U68389; AAB19199.1; -.
DR TRANSFAC; T03410; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1187 MW; 278B51F37B11F40B CRC64;

Query Match 48.9%; Score 22; DB 13; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WHF 6
Db 5 WHY 7

RESULT 7
Q9B4X0
ID Q9B4X0 PRELIMINARY; PRT; 10 AA.
AC Q9B4X0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit 1 (Fragment).
GN COI.
OS Notoththalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC Notoththalmus.
OX NCBI_TaxID=8316;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21175761; PubMed=11277635;
RA Weisrock D.W., Macey J.R., Ugurtas I.H., Larson A., Papenfuss T.J.;
RT "Molecular Phylogenetics and Historical Biogeography among
Salamanders of the 'True' Salamander Clade: Rapid Branching of
Numerous Highly Divergent Lineages in Mertensiella luschni Associated
with the Rise of Anatolia.";
RL Mol. Phylogenet. Evol. 18:434-448(2001).
DR EMBL: AF296616; AAK30305.1; -.
DR GO: GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1298 MW; 03D380C733640050 CRC64;

Query Match 48.9%; Score 22; DB 8; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VRWHF 6
Db 4 IRWLF 8

```

```

RESULT 8
Q958L2
ID Q958L2 PRELIMINARY; PRT; 10 AA.
AC Q958L2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Rana temporaria (European common frog).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21184280; PubMed=11286498;
RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
RA Jennings M., Larson A.;
RT "Molecular Phylogenetics of Western North American Frogs of the Rana
boylii Species Group.";
RL Mol. Phylogenet. Evol. 19:131-143(2001).
DR EMBL: AF314018; AAK56874.1; -.
DR GO: GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1354 MW; COD380C9D36411A9 CRC64;

Query Match 48.9%; Score 22; DB 8; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WVRWHF 6
Db 3 FTRWFF 8

RESULT 9
Q958L8
ID Q958L8 PRELIMINARY; PRT; 10 AA.
AC Q958L8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Rana catesbeiana (Bull frog).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21184280; PubMed=11286498;
RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
RA Jennings M., Larson A.;
RT "Molecular Phylogenetics of Western North American Frogs of the Rana
boylii Species Group.";
RL Mol. Phylogenet. Evol. 19:131-143(2001).
DR EMBL: AF314016; AAK56868.1; -.
DR GO: GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1354 MW; COD380C9D36411A9 CRC64;

Query Match 48.9%; Score 22; DB 8; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WVRWHF 6
Db 3 FTRWFF 8

```

Db 3 FTRWFF 8

RESULT 10

Q958K6 PRELIMINARY; PRT; 10 AA.
 ID Q958K6
 AC Q958K6
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Rana pretiosa.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
 OX NCBI_TaxID=69834;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21184280; PubMed=11286498;
 RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
 RA Jennings M., Larson A.;
 RT "Molecular Phylogenetics of Western North American Frogs of the Rana
 RT boylii Species Group."
 RL Mol. Phylogenet. Evol. 19:131-143 (2001).
 DR EMBL; AF314020; AAK56886.1; -
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON TER 10 10
 SQ SEQUENCE 10 AA; 1354 MW; COD380C9D36411A9 CRC64;

Query Match 48.9%; Score 22; DB 8; Length 10;
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWFF 6
 : |||
 Db 3 FTRWFF 8

RESULT 11

Q958K0 PRELIMINARY; PRT; 10 AA.
 ID Q958K0
 AC Q958K0
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Rana cascadae.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
 OX NCBI_TaxID=160497;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21184280; PubMed=11286498;
 RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
 RA Jennings M., Larson A.;
 RT "Molecular Phylogenetics of Western North American Frogs of the Rana
 RT boylii Species Group."
 RL Mol. Phylogenet. Evol. 19:131-143 (2001).
 DR EMBL; AF314022; AAK56886.1; -
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON TER 10 10
 SQ SEQUENCE 10 AA; 1354 MW; COD380C9D36411A9 CRC64;

Query Match 48.9%; Score 22; DB 8; Length 10;
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWFF 6

Db 3 FTRWFF 8
 : |||

RESULT 12

Q958L5 PRELIMINARY; PRT; 10 AA.
 ID Q958L5
 AC Q958L5
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Rana sylvatica (Wood frog).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
 OX NCBI_TaxID=45438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21184280; PubMed=11286498;
 RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
 RA Jennings M., Larson A.;
 RT "Molecular Phylogenetics of Western North American Frogs of the Rana
 RT boylii Species Group."
 RL Mol. Phylogenet. Evol. 19:131-143 (2001).
 DR EMBL; AF314017; AAK56871.1; -
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON TER 10 10
 SQ SEQUENCE 10 AA; 1354 MW; COD380C9D36411A9 CRC64;

Query Match 48.9%; Score 22; DB 8; Length 10;
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVRWFF 6
 : |||
 Db 3 FTRWFF 8

RESULT 13

Q958K3 PRELIMINARY; PRT; 10 AA.
 ID Q958K3
 AC Q958K3
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Rana aurora (Red-legged frog).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
 OX NCBI_TaxID=160496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21184280; PubMed=11286498;
 RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
 RA Jennings M., Larson A.;
 RT "Molecular Phylogenetics of Western North American Frogs of the Rana
 RT boylii Species Group."
 RL Mol. Phylogenet. Evol. 19:131-143 (2001).
 DR EMBL; AF314021; AAK56883.1; -
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON TER 10 10
 SQ SEQUENCE 10 AA; 1354 MW; COD380C9D36411A9 CRC64;

Query Match 48.9%; Score 22; DB 8; Length 10;
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 WVRWHF 6
   : |||
Db 3 FTRWFF 8

RESULT 14
Q94NH4
ID Q94NH4 PRELIMINARY; PRT; 10 AA.
AC Q94NH4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Rana muscosa.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=160500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21184280; PubMed=11286498;
RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
RA Jennings M., Larson A.;
RT "Molecular Phylogenetics of Western North American Frogs of the Rana
RT boylei Species Group.";
RL Mol. Phylogenet. Evol. 19:131-143 (2001).
DR EMBL; AF314023; AAK56889.1; -
DR EMBL; AF314024; AAK56892.1; -
DR EMBL; AF314025; AAK56895.1; -
DR EMBL; AF314027; AAK56901.1; -
DR EMBL; AF314028; AAK56904.1; -
DR EMBL; AF314029; AAK56907.1; -
DR EMBL; AF314030; AAK56910.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 10 10
SQ SEQUENCE 10 AA; 1354 MW; C0D380C9D36411A9 CRC64;

Query Match 48.9%; Score 22; DB 8; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WVRWHF 6
   : |||
Db 3 FTRWFF 8

RESULT 15
Q94VD2
ID Q94VD2 PRELIMINARY; PRT; 10 AA.
AC Q94VD2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus panoptes panoptes.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169849;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0 (2001).
DR EMBL; AF407516; AAL10102.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 10 10
SQ SEQUENCE 10 AA; 1299 MW; 5DEE80D4136411A7 CRC64;

```

```

Query Match 48.9%; Score 22; DB 8; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 3 RWHF 6
   : |||
Db 5 RWRFF 8

```

Search completed: June 10, 2004, 11:18:43
 Job time : 32.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:51:30 ; Search time 46.3333 Seconds
(without alignments)
36.589 Million cell updates/sec

Title: US-09-912-414-9

Perfect score: 31

Sequence: 1 WXXWXP 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 447529

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesep 29Jan04:*

- 1: Genesep1980s:*
- 2: Genesep1990s:*
- 3: Genesep2000s:*
- 4: Genesep2001s:*
- 5: Genesep2002s:*
- 6: Genesep2003as:*
- 7: Genesep2003bs:*
- 8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	90.3	15	2	AAY30351 Epitope d
2	28	90.3	15	5	AAB19239 Streptoco
3	27	87.1	15	5	AAB26751 Fibrin bi
4	27	87.1	15	5	AAB26733 Fibrin bi
5	26	83.9	6	3	AAB01508 Peptide w
6	26	83.9	6	3	AAB01505 Peptide w
7	26	83.9	6	3	AAB01506 Peptide w
8	26	83.9	6	6	ABR45426 Peptide w
9	26	83.9	6	6	ABR45594 Peptide w
10	26	83.9	6	6	ABR45314 Peptide w
11	26	83.9	6	6	ABR45481 Peptide w
12	26	83.9	6	6	ABR45369 Peptide w
13	26	83.9	6	6	ABR45425 Peptide w
14	26	83.9	6	6	ABR45593 Peptide w
15	26	83.9	6	6	ABR45313 Peptide w
16	26	83.9	6	6	ABR45370 Peptide w
17	26	83.9	6	6	ABR45482 Peptide w
18	26	83.9	9	5	AAE26775 Fibrin bi
19	26	83.9	15	3	AAE26759 Fibrin bi
20	26	83.9	15	5	AAE26759 Fibrin bi
21	26	83.9	15	5	AAU86245 Oestrogen
22	26	83.9	15	5	AAU86245 Oestrogen
23	25	80.6	6	2	AAR57391 Peptide f
24	25	80.6	6	3	AAB01499 Peptide w
25	25	80.6	6	3	AAB01492 Peptide w

26	25	80.6	6	3	AAB01497	Aab01497 Peptide w
27	25	80.6	6	6	ABR44865	ABR44865 Staphyloc
28	25	80.6	6	6	ABR45592	ABR45592 Staphyloc
29	25	80.6	6	6	ABR45311	ABR45311 Staphyloc
30	25	80.6	6	6	ABR45480	ABR45480 Staphyloc
31	25	80.6	6	6	ABR45591	ABR45591 Staphyloc
32	25	80.6	6	6	ABR45312	ABR45312 Staphyloc
33	25	80.6	6	6	ABR45424	ABR45424 Staphyloc
34	25	80.6	6	6	ABR45367	ABR45367 Staphyloc
35	25	80.6	6	6	ABR45423	ABR45423 Staphyloc
36	25	80.6	6	6	ABR44866	ABR44866 Staphyloc
37	25	80.6	6	6	ABR45538	ABR45538 Staphyloc
38	25	80.6	6	6	ABR45479	ABR45479 Staphyloc
39	25	80.6	6	6	ABR45537	ABR45537 Staphyloc
40	25	80.6	6	6	ABR45368	ABR45368 Staphyloc
41	25	80.6	7	4	AAM45777	Aam45777 H11 bindi
42	25	80.6	9	3	AAB01498	Aab01498 Peptide w
43	25	80.6	11	3	AAB20714	Aab20714 PolymERIC
44	25	80.6	13	2	AAM38112	Aam38112 Dystrophin
45	25	80.6	13	7	ADB49262	ADB49262 Dystrophin

ALIGNMENTS

RESULT 1
AAY30351
ID AAY30351 standard; peptide; 15 AA.
XX
AC AAY30351;
XX
DT 09-NOV-1999 (first entry)
XX
DE Epitope derived from pneumococcal surface adhesion A protein.
XX
KW Pneumococcal surface adhesion A protein; PaaA; monoclonal antibody;
KW vaccine; Streptococcus pneumoniae infection.
XX
OS Streptococcus pneumoniae.
XX
FN WO9945121-A1.
XX
PD 10-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US004326.
XX
PR 02-MAR-1998; 98US-0076565P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Carlone GM, Ades EW, Sampson JS, Tharpe JA, Zeiler JL;
PI Westerink MAJ;
XX
DR WPI; 1999-540849/45.
XX
PT New peptides corresponding to Streptococcus pneumoniae paaA, used for
PT treating or preventing Streptococcus pneumoniae infection in a subject.
XX
FS Claim 6; Page 43; 58pp; English.
XX
CC AAY30351-54 represent immunogenic peptides which are derived from a
CC pneumococcal surface adhesion A protein (PaaA). The specification
CC describes monoclonal antibodies which bind epitopes of the PaaA protein
CC (e.g present sequence). The peptides can be used in vaccines to prevent
CC Streptococcus pneumoniae infections. The antibodies of the invention can
CC also be used to detect S. pneumoniae in a sample or individual
XX
SQ Sequence 15 AA;

Query Match 90.3%; Score 28; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWXP 6
Db 7 WTAWAF 12

RESULT 2
AAE19239
ID AAE19239 standard; peptide; 15 AA.
XX AC AAE19239;
XX DT 21-MAY-2002 (first entry)
XX DE Streptococcus pneumoniae PsaA immunogenic peptide #1.
XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX KW pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX OS Streptococcus pneumoniae.
XX PN WO200204497-A2.
XX XX 17-JAN-2002.
XX PF 10-JUL-2001; 2001WO-US021626.
XX PR 10-JUL-2000; 2000US-00613092.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX DR WPI; 2002-195762/25.
XX XX
XX PT New multiple antigenic peptide for immunizing against streptococcal
XX PT infections, binds to monoclonal antibody obtained in response to
XX PT immunizing an animal with pneumococcal surface adhesion protein A or its
XX XX fragment.
XX PS Claim 2; Page 56; 86pp; English.
XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
XX CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX CC obtained in response to immunising an animal with pneumococcal surface
XX CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX CC protective immunity against S. pneumoniae infection in a subject. The
XX CC present sequence is Streptococcus pneumoniae PsaA immunogenic peptide
XX SQ Sequence 15 AA;
Query Match 90.3%; Score 28; DB 5; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWXP 6
Db 7 WTAWAF 12

RESULT 3
AAE26751
ID AAE26751 standard; peptide; 9 AA.
XX AC AAE26751;
XX DT 13-DEC-2002 (first entry)
XX DE Fibrin binding loop #3.
XX KW Fibrin binding peptide; thrombosis; pulmonary embolism; atherosclerosis;
XX KW myocardial infarct; ischaemia; imaging; rheumatoid arthritis; vasotropic;
XX KW anaemia; hypoxia; tumour; diabetic retinopathy; autoimmune disorder;
XX KW inflammatory disorder; angiogenesis; stroke; cerebroprotective.

Unidentified.
WO200255544-A2.
18-JUL-2002.
21-DEC-2001; 2001WO-US049534.
23-DEC-2000; 2000US-00747403.
(DYAX-) DYAX CORP.
Wescott CR, Beltzer JP, Sato AK;
WPI; 2002-666875/71.
Novel synthetic fibrin-binding moiety, useful for detecting, imaging or
localizing fibrin-containing clots by magnetic resonance imaging,
radioimaging and for treating diseases involving thrombus formation e.g.
stroke.
Claim 4; Page 55; 89pp; English.
The invention relates to a synthetic fibrin binding group having affinity
for fibrin. The invention is useful for detecting fibrin in a mammalian
subject which involves (a) detectably labelling the binding group; (b)
administering to the subject the labelled polypeptide, and (c) detecting
the labelled polypeptide in the subject. The invention is useful for
treating a disease involving thrombus formation eg. deep-vein thrombosis,
pulmonary embolism, cardiogenic thrombosis, atherosclerosis, myocardial
infarct, reperfusion ischaemia or stroke. The binding moieties are useful
for detection, imaging and localisation of fibrin-containing clots by
magnetic resonance imaging, radioimaging and other imaging methods and
are also useful in the diagnosis and treatment of coronary conditions
where fibrin plays a role. The fibrin binding moieties are useful for
detecting and diagnosing numerous pathophysiological conditions in which fibrin plays
a role eg. peritoneal adhesions which often occur after surgery or
inflammatory and neoplastic processes and are comprised of a fibrin
network, fibroblasts, macrophages and new blood vessels; rheumatoid
arthritis, lupus or septic arthritis which often have bits of fibrin
containing tissues called rice bodies in the synovial fluid of their
joints; thrombocytopenic purpura, a type of anaemia in which deposits in
arterioles causes turbulent blood flow resulting in stress and
destruction of red blood cells. The fibrin specific agents can also be
used to detect hypoxia or ischaemia of heart, kidney, liver, lung, brain
or other organs, as well as the detection of tumours, diabetic
retinopathy, early or high-risk atherosclerosis and other autoimmune and
inflammatory disorders. Fibrin specific agents also could provide both
direct or surrogate markers of disease models in which hypoxia and
angiogenesis are expected to play a role. The invention is also useful
for screening molecular libraries. The present sequence is a fibrin
binding loop
Sequence 9 AA;
Query Match 87.1%; Score 27; DB 5; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 WXXWXP 6
Db 3 WESWTF 8

RESULT 4
AAE26733
ID AAE26733 standard; peptide; 15 AA.
XX AC AAE26733;
XX DT 13-DEC-2002 (first entry)
XX XX

DE Fibrin binding peptide #4.
 XX Fibrin binding peptide; thrombosis; pulmonary embolism; atherosclerosis;
 KW myocardial infarct; ischaemia; imaging; rheumatoid arthritis; vasotropic;
 KW anaemia; hypoxia; tumour; diabetic retinopathy; autoimmune disorder;
 KW inflammatory disorder; angiogenesis; stroke; cerebroprotective.
 XX Unidentified.
 OS
 FN WO200255544-A2.
 XX 18-JUL-2002.
 XX
 XX 21-DEC-2001; 2001WO-US049534.
 XX
 XX 23-DEC-2000; 2000US-00747403.
 XX (DYAX-) DYAX CORP.
 XX
 XX Wescott CR, Beltzer JP, Sato AK;
 XX
 XX WPI; 2002-666875/71.
 XX
 XX Novel synthetic fibrin-binding moiety, useful for detecting, imaging or
 PT localizing fibrin-containing clots by magnetic resonance imaging,
 PT radioimaging and for treating diseases involving thrombus formation e.g.
 PT stroke.
 XX
 XX Claim 10; Page 57; 89pp; English.

XX The invention relates to a synthetic fibrin binding group having affinity
 CC for fibrin. The invention is useful for detecting fibrin in a mammalian
 CC subject which involves (a) detectably labelling the binding group; (b)
 CC administering to the subject the labelled polypeptide, and (c) detecting
 CC the labelled polypeptide in the subject. The invention is useful for
 CC treating a disease involving thrombus formation eg. deep-vein thrombosis,
 CC pulmonary embolism, cardiogenic thrombosis, atherosclerosis, myocardial
 CC infarct, reperfusion ischaemia or stroke. The binding moieties are useful
 CC for detection, imaging and localisation of fibrin-containing clots by
 CC magnetic resonance imaging, radioimaging and other imaging methods and
 CC are also useful in the diagnosis and treatment of coronary conditions
 CC where fibrin plays a role. The fibrin binding moieties are useful for
 CC detecting and diagnosing numerous pathophysiological in which fibrin plays
 CC a role eg. peritoneal adhesions which often occur after surgery or
 CC inflammatory and neoplastic processes and are comprised of a fibrin
 CC network, fibroblasts, macrophages and new blood vessels; rheumatoid
 CC arthritis, lupus or septic arthritis which often have bits of fibrin
 CC containing tissues called rice bodies in the synovial fluid of their
 CC joints; thrombocytopenic purpura, a type of anaemia in which deposits in
 CC arterioles causes turbulent blood flow resulting in stress and
 CC destruction of red blood cells. The fibrin specific agents can also be
 CC used to detect hypoxia or ischaemia of heart, kidney, liver, lung, brain
 CC or other organs, as well as the detection of tumours, diabetic
 CC retinopathy, early or high-risk atherosclerosis and other autoimmune and
 CC inflammatory disorders. Fibrin specific agents also could provide both
 CC direct or surrogate markers of disease models in which hypoxia and
 CC angiogenesis are expected to play a role. The invention is also useful
 CC for screening molecular libraries. The present sequence is a fibrin
 CC binding peptide

XX Sequence 15 AA;

Query Match 87.1%; Score 27; DB 5; Length 15;
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 | | |
 Db 6 WESWTF 11

RESULT 5
 AAB01508

ID AAB01508 standard; peptide; 6 AA.
 XX
 AC AAB01508;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
 XX
 KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
 KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
 KW restenosis.
 XX
 OS Synthetic.
 XX
 XX WO2000044771-A1.
 FN
 XX 03-AUG-2000.
 PD
 XX 26-JAN-2000; 2000WO-GB0000227.
 PF
 XX 26-JAN-1999; 99GB-00001710.
 PR
 XX (PROL-) PROLIFIX LTD.
 PA
 XX Mueller R, Kontermann RE, Montigiani S;
 PI
 XX WPI; 2000-532806/48.
 DR
 XX Peptides binding to the DNA binding domain of transcription factor E2F
 PT and inhibiting cell cycle progression, useful for the treatment of
 PT cancer.
 XX
 XX Example; Page 26; 42pp; English.
 XX
 CC Peptides which bind to the DNA binding domain of transcription factor E2F
 CC and inhibit cell cycle progression may be useful as research agents to
 CC investigate the interaction between E2F and DP-1, or the activation of
 CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
 CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
 CC treatment of cancer or other proliferative disorders such as psoriasis
 CC and restenosis
 CC
 XX Sequence 6 AA;
 SQ
 Query Match 83.9%; Score 26; DB 3; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWKF 6
 | | |
 Db 1 WVRWAF 6
 RESULT 6
 AAB01505
 ID AAB01505 standard; peptide; 6 AA.
 XX
 AC AAB01505;
 XX
 XX 08-NOV-2000 (first entry)
 DT
 XX
 DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
 XX
 KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
 KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
 KW restenosis.
 XX
 OS Synthetic.
 XX
 XX WO2000044771-A1.
 FN
 XX 03-AUG-2000.
 PD
 XX

PF 26-JAN-2000; 2000WO-GB0000227.
 XX
 PR 26-JAN-1999; 99GB-00001710.
 XX
 PA (PROL-) PROLIFIX LTD.
 XX
 XX Mueller R, Kontermann RE, Montigiani S;
 PI WPI; 2000-532806/48.
 XX
 DR Peptides binding to the DNA binding domain of transcription factor E2F
 XX and inhibiting cell cycle progression, useful for the treatment of
 PT cancer.
 PT
 XX Example; Page 26; 42pp; English.
 XX
 CC Peptides which bind to the DNA binding domain of transcription factor E2F
 CC and inhibit cell cycle progression may be useful as research agents to
 CC investigate the interaction between E2F and DP-1, or the activation of
 CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
 CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
 CC treatment of cancer or other proliferative disorders such as psoriasis
 CC and restenosis
 XX
 XX Sequence 6 AA;
 SQ

Query Match 83.9%; Score 26; DB 3; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 1 WARWHF 6

RESULT 7
 ID AAB01506 standard; peptide; 6 AA.
 AC AAB01506;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
 XX
 KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
 KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
 KW restenosis.
 XX
 OS Synthetic.
 XX
 PN WO200044771-A1.
 XX
 PD 03-AUG-2000.
 XX
 PF 26-JAN-2000; 2000WO-GB0000227.
 XX
 PR 26-JAN-1999; 99GB-00001710.
 XX
 XX (PROL-) PROLIFIX LTD.
 XX
 PI Mueller R, Kontermann RE, Montigiani S;
 XX WPI; 2000-532806/48.
 DR
 DR Peptides binding to the DNA binding domain of transcription factor E2F
 PT and inhibiting cell cycle progression, useful for the treatment of
 PT cancer.
 XX
 XX Example; Page 26; 42pp; English.
 PS
 PS Peptides which bind to the DNA binding domain of transcription factor E2F
 CC and inhibit cell cycle progression may be useful as research agents to

CC investigate the interaction between E2F and DP-1, or the activation of
 CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
 CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
 CC treatment of cancer or other proliferative disorders such as psoriasis
 CC and restenosis
 XX
 XX Sequence 6 AA;
 SQ

Query Match 83.9%; Score 26; DB 3; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 Db 1 WVAWHF 6

RESULT 8
 ID ABR45426 standard; peptide; 6 AA.
 XX
 AC ABR45426;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 DE Staphylococcus aureus CHIPS-related peptide #616.
 XX
 KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formulated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX
 PN WO2003006048-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2001; 2001WO-EP008004.
 XX
 PR 11-JUL-2001; 2001WO-EP008004.
 XX
 XX (JARI-) JARI PHARM BV.
 XX
 PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX
 DR WPI; 2003-256333/25.
 XX
 PT Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases.
 XX
 XX Disclosure; Page 12; 89pp; English.
 PS
 PS The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formulated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, disease of the central nervous system, gastrointestinal
 CC diseases, skin diseases, genitourinary diseases, joint diseases,
 CC respiratory diseases and HIV infection
 XX
 XX Sequence 6 AA;
 SQ

Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 DB 1 WTFWLF 6

RESULT 9

ABR45594
 ID ABR45594 standard; peptide; 6 AA.

AC ABR45594;

DT 10-JUN-2003 (first entry)

DE Staphylococcus aureus CHIPS-related peptide #784.

XX CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formulated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nontropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.

OS Staphylococcus aureus.

OS Synthetic.

PN WO2003006048-A1.

PD 23-JAN-2003.

PF 11-JUL-2001; 2001WO-EP008004.

PR 11-JUL-2001; 2001WO-EP008004.

PA (JARI-) JARI PHARM BV.

PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;

PS WPI; 2003-256333/25.

XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases.

PS Disclosure; Page 13; 89pp; English.

XX The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formulated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, disease of the central nervous system, gastrointestinal
 CC diseases, skin diseases, genitourinary diseases, joint diseases,
 CC respiratory diseases and HIV infection

XX Sequence 6 AA;

Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 DB 1 WTFWYF 6

RESULT 10

ABR45314
 ID ABR45314 standard; peptide; 6 AA.

XX ABR45314;

DT 10-JUN-2003 (first entry)

DE Staphylococcus aureus CHIPS-related peptide #504.

XX CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formulated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nontropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.

OS Staphylococcus aureus.

OS Synthetic.

PN WO2003006048-A1.

PD 23-JAN-2003.

PF 11-JUL-2001; 2001WO-EP008004.

PR 11-JUL-2001; 2001WO-EP008004.

PA (JARI-) JARI PHARM BV.

PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;

PS WPI; 2003-256333/25.

XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases.

PS Disclosure; Page 12; 89pp; English.

XX The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formulated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, disease of the central nervous system, gastrointestinal
 CC diseases, skin diseases, genitourinary diseases, joint diseases,
 CC respiratory diseases and HIV infection

XX Sequence 6 AA;

Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWKF 6
 DB 1 WTFWFF 6

RESULT 11

ABR45481
 ID ABR45481 standard; peptide; 6 AA.

XX ABR45481;

DT 10-JUN-2003 (first entry)

XX Staphylococcus aureus CHIPS-related peptide #671.
 DE CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 XX formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX Staphylococcus aureus.
 OS Synthetic.
 XX WO2003006048-A1.
 PN 23-JAN-2003.
 XX 11-JUL-2001; 2001WO-EP008004.
 PF (JARI-) JARI PHARM BV.
 XX Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX WPI; 2003-256333/25.
 DR Combination of peptides derived from chemotaxis inhibiting protein from
 XX Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases.
 XX Disclosure; Page 13; 89pp; English.
 PS The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, disease of the central nervous system, gastrointestinal
 CC diseases, skin diseases, genitourinary diseases, joint diseases,
 CC respiratory diseases and HIV infection
 XX Sequence 6 AA;
 SQ
 Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 DB 1 WSWFVF 6
 RESULT 12
 ABR45369
 ID ABR45369 standard; peptide; 6 AA.
 XX ABR45369;
 AC 10-JUN-2003 (first entry)
 DT Staphylococcus aureus CHIPS-related peptide #559.
 DE CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 XX

KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX Staphylococcus aureus.
 OS Synthetic.
 XX WO2003006048-A1.
 PN 23-JAN-2003.
 XX 11-JUL-2001; 2001WO-EP008004.
 PF (JARI-) JARI PHARM BV.
 XX Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX WPI; 2003-256333/25.
 DR Combination of peptides derived from chemotaxis inhibiting protein from
 XX Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases.
 XX Disclosure; Page 12; 89pp; English.
 PS The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, disease of the central nervous system, gastrointestinal
 CC diseases, skin diseases, genitourinary diseases, joint diseases,
 CC respiratory diseases and HIV infection
 XX Sequence 6 AA;
 SQ
 Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXF 6
 DB 1 WSWFVF 6
 RESULT 13
 ABR45425
 ID ABR45425 standard; peptide; 6 AA.
 XX ABR45425;
 AC 10-JUN-2003 (first entry)
 DT Staphylococcus aureus CHIPS-related peptide #615.
 DE CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX Staphylococcus aureus.
 OS Synthetic.
 XX WO2003006048-A1.
 PN

PD 23-JAN-2003.
 XX
 PF 11-JUL-2001; 2001WO-EP008004.
 XX
 PR 11-JUL-2001; 2001WO-EP008004.
 XX
 PA (JARI-) JARI PHARM BV.
 XX
 PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PF Van Strijp JAG;
 XX
 DR WPI; 2003-256333/25.
 XX
 PS Combination of peptides derived from chemotaxis inhibiting protein from
 XX Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 XX prophylaxis and treatment of inflammation, cardiovascular, skin and
 XX kidney diseases.
 XX
 PS Disclosure; Page 12; 89pp; English.
 XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, skin diseases, genitourinary system, gastrointestinal
 CC respiratory diseases and HIV infection
 XX
 SQ Sequence 6 AA;
 Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 DB 1 WSWWLP 6
 RESULT 14
 ABR45593
 ID ABR45593 standard; peptide; 6 AA.
 AC ABR45593;
 XX
 DT 10-JUN-2003 (first entry)
 DE Staphylococcus aureus CHIPS-related peptide #783.
 KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nontropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX
 FN WO2003006048-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2001; 2001WO-EP008004.
 XX
 PR 11-JUL-2001; 2001WO-EP008004.
 XX
 PA (JARI-) JARI PHARM BV.
 XX
 PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PF Van Strijp JAG;
 XX
 DR WPI; 2003-256333/25.
 XX
 PS Combination of peptides derived from chemotaxis inhibiting protein from
 XX Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 XX prophylaxis and treatment of inflammation, cardiovascular, skin and
 XX kidney diseases.
 XX
 PS Disclosure; Page 12; 89pp; English.
 XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, skin diseases, genitourinary system, gastrointestinal
 CC respiratory diseases and HIV infection
 XX
 SQ Sequence 6 AA;
 Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 DB 1 WSWWLP 6
 RESULT 14
 ABR45593
 ID ABR45593 standard; peptide; 6 AA.
 AC ABR45593;
 XX
 DT 10-JUN-2003 (first entry)
 DE Staphylococcus aureus CHIPS-related peptide #783.
 KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nontropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX
 FN WO2003006048-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2001; 2001WO-EP008004.
 XX
 PR 11-JUL-2001; 2001WO-EP008004.
 XX
 PA (JARI-) JARI PHARM BV.
 XX
 PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PF Van Strijp JAG;
 XX
 DR WPI; 2003-256333/25.
 XX
 PS Combination of peptides derived from chemotaxis inhibiting protein from
 XX Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 XX prophylaxis and treatment of inflammation, cardiovascular, skin and
 XX kidney diseases.
 XX
 PS Disclosure; Page 12; 89pp; English.
 XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, skin diseases, genitourinary system, gastrointestinal
 CC respiratory diseases and HIV infection
 XX
 SQ Sequence 6 AA;
 Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 DB 1 WSWWLP 6
 RESULT 15
 ABR45313
 ID ABR45313 standard; peptide; 6 AA.
 AC ABR45313;
 XX
 DT 10-JUN-2003 (first entry)
 DE Staphylococcus aureus CHIPS-related peptide #503.
 KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nontropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX
 FN WO2003006048-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2001; 2001WO-EP008004.
 XX
 PR 11-JUL-2001; 2001WO-EP008004.
 XX
 PA (JARI-) JARI PHARM BV.
 XX
 PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PF Van Strijp JAG;
 XX
 DR WPI; 2003-256333/25.
 XX
 PS Combination of peptides derived from chemotaxis inhibiting protein from
 XX Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 XX prophylaxis and treatment of inflammation, cardiovascular, skin and
 XX kidney diseases.
 XX
 PS Disclosure; Page 13; 89pp; English.
 XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, skin diseases, genitourinary system, gastrointestinal
 CC respiratory diseases and HIV infection
 XX
 SQ Sequence 6 AA;
 Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 DB 1 WSWWLP 6

PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX
 DR WPI; 2003-256333/25.
 XX
 PS Combination of peptides derived from chemotaxis inhibiting protein from
 XX Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 XX prophylaxis and treatment of inflammation, cardiovascular, skin and
 XX kidney diseases.
 XX
 PS Disclosure; Page 13; 89pp; English.
 XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, skin diseases, genitourinary system, gastrointestinal
 CC respiratory diseases and HIV infection
 XX
 SQ Sequence 6 AA;
 Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 DB 1 WSWWLP 6
 RESULT 15
 ABR45313
 ID ABR45313 standard; peptide; 6 AA.
 AC ABR45313;
 XX
 DT 10-JUN-2003 (first entry)
 DE Staphylococcus aureus CHIPS-related peptide #503.
 KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nontropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX
 FN WO2003006048-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2001; 2001WO-EP008004.
 XX
 PR 11-JUL-2001; 2001WO-EP008004.
 XX
 PA (JARI-) JARI PHARM BV.
 XX
 PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PF Van Strijp JAG;
 XX
 DR WPI; 2003-256333/25.
 XX
 PS Combination of peptides derived from chemotaxis inhibiting protein from
 XX Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 XX prophylaxis and treatment of inflammation, cardiovascular, skin and
 XX kidney diseases.
 XX
 PS Disclosure; Page 13; 89pp; English.
 XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, skin diseases, genitourinary system, gastrointestinal
 CC respiratory diseases and HIV infection
 XX
 SQ Sequence 6 AA;
 Query Match 83.9%; Score 26; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WXXWXP 6
 DB 1 WSWWLP 6

PT Kidney diseases.

XX
PS Disclosure; Page 12; 89pp; English.

XX
CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
CC Staphylococcus aureus. The peptide fragments are useful in the
CC prophylaxis or treatment of diseases or disorders involving the C5a-
CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
CC monocytes and endothelial cells or involving acute or chronic
CC inflammation reactions. The diseases or disorders include cardiovascular
CC diseases, disease of the central nervous system, gastrointestinal
CC diseases, skin diseases, genitourinary diseases, joint diseases,
CC respiratory diseases and HIV infection

XX
SQ Sequence 6 AA;

Query Match 83.9%; Score 26; DB 6; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWYF 6

Db 1 WSWFF 6

Search completed: June 10, 2004, 11:06:18
Job time : 47.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:52:15 ; Search time 11.333 Seconds
(without alignments)
50.925 Million cell updates/sec

Title: US-09-912-414-9

Perfect score: 31

Sequence: 1 WXXWF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 2522

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 78:*

2: PIR1:*

3: PIR2:*

4: PIR3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	67.7	9	2 A43848	cell surface adhes
2	20	64.5	10	2 F49033	T-cell receptor ga
3	20	64.5	12	2 PH1324	Ig heavy chain DJ
4	20	64.5	12	2 PH1308	Ig heavy chain DJ
5	20	64.5	13	2 S61798	T-cell-specific tr
6	20	64.5	14	2 PH1322	Ig heavy chain DJ
7	17	54.8	13	2 S23372	T-cell receptor al
8	17	54.8	13	2 B25448	Ig kappa-1 chain,
9	17	54.8	13	2 B26406	Ig kappa chain J r
10	17	54.8	13	2 A47630	Ig kappa chain J r
11	16	51.6	8	2 T13818	cytochrome oxidase
12	16	51.6	10	2 T17054	cytochrome-c oxida
13	16	51.6	10	2 T13976	cytochrome-c oxida
14	16	51.6	10	2 T17057	cytochrome-c oxida
15	16	51.6	10	2 T12303	cytochrome-c oxida
16	16	51.6	10	2 T14019	cytochrome-c oxida
17	16	51.6	10	2 T17060	cytochrome-c oxida
18	16	51.6	10	2 T17063	cytochrome-c oxida
19	16	51.6	10	2 T12335	cytochrome-c oxida
20	16	51.6	10	2 T14043	cytochrome-c oxida
21	16	51.6	10	2 T14054	cytochrome-c oxida
22	16	51.6	10	2 T17066	cytochrome-c oxida
23	16	51.6	10	2 T17069	cytochrome-c oxida
24	16	51.6	10	2 T12308	cytochrome-c oxida
25	16	51.6	10	2 T17072	cytochrome-c oxida
26	16	51.6	10	2 T12312	cytochrome-c oxida
27	16	51.6	10	2 T12329	cytochrome-c oxida
28	16	51.6	10	2 T12316	cytochrome-c oxida
29	16	51.6	10	2 T12321	cytochrome-c oxida

```

30      16      51.6      10      2      T14215      cytochrome-c oxida
31      16      51.6      10      2      T14223      cytochrome-c oxida
32      16      51.6      10      2      T14219      cytochrome-c oxida
33      16      51.6      12      2      A29189      phospholipase A2 (
34      16      51.6      14      2      PT0077      proteochondotin c
35      16      51.6      15      2      PA0099      phenotypic variati
36      15      48.4      9      2      S56004      glucan 1,3-beta-gl
37      15      48.4      9      2      S24159      leukocyte elastase
38      14      45.2      9      2      S07241      litorin - Rohde's
39      14      45.2      10      1      RHPGG      gonadoliberin - pi
40      14      45.2      10      1      RSHHG      gonadoliberin - sh
41      14      45.2      10      1      RHAQI      gonadoliberin I -
42      14      45.2      10      2      A21114      gonadoliberin - ch
43      14      45.2      11      2      S68649      spermadhesin AQN-3
44      14      45.2      15      2      PH1365      Ig heavy chain DJ
45      13      41.9      9      2      S07205      litorin 2-Glu - Au

```

ALIGNMENTS

RESULT 1

A43848
cell surface adhesin for heparan sulfate, 66K - Staphylococcus aureus (fragment)
C:Species: Staphylococcus aureus
C>Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 24-Feb-1995
C:Accession: A43848
R:Liang, O.D.; Ascencio, F.; Fransson, L.A.; Wadstrom, T.
Infect. Immun. 60, 899-906, 1992
A:Title: Binding of heparan sulfate to Staphylococcus aureus.
A:Reference number: A43848; MUID:92176005; PMID:1541563
A:Accession: A43848
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <LIA>
A>Note: sequence extracted from NCBI backbone (NCBIP:85442)

Query Match 67.7%; Score 21; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXW 4
DB 2 WTGW 5

RESULT 2

F49033
T-cell receptor gamma chain V-D-J region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Dec-1993 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: F49033
R:Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.
Eur. J. Immunol. 21, 2999-3007, 1991
A:Title: Functionally distinct subsets of human gamma/delta T cells.
A:Reference number: A49033; MUID:92083926; PMID:1684157
A:Accession: F49033
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-10 <MOR>
A:Cross-references: GB:S72605; PIDN:AAB20632.1; PID:G240701
A>Note: sequence extracted from NCBI backbone (NCBIN:72605, NCBIP:72606)
C:Keywords: T-cell receptor

Query Match 64.5%; Score 20; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXW 4
DB 4 WERW 7

RESULT 3

PHI324
IG heavy chain DJ region (clone C510-100) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI324
R:Wasserman, R.; Galli, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
A:Reference number: PHI324; MUID:93094761; PMID:1460419
A:Accession: PHI324
A:Molecule type: DNA
A:Residues: 1-12 <WAS>
C:Keywords: heterotetramer; immunoglobulin

Query Match 64.5%; Score 20; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 6.5e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
|
|
Db 5 WYTW 8

RESULT 4

PHI308
IG heavy chain DJ region (clone C731-94) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI308
R:Wasserman, R.; Galli, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
A:Reference number: PHI308; MUID:93094761; PMID:1460419
A:Accession: PHI308
A:Molecule type: DNA
A:Residues: 1-12 <WAS>
C:Keywords: heterotetramer; immunoglobulin

Query Match 64.5%; Score 20; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 6.5e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
|
|
Db 7 WGQW 10

RESULT 5

S61798
T-cell-specific transcription factor 1 splice form G - human (fragment)
N:Alternate names: transcription factor TCF-1G
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 24-Jul-1998
C:Accession: S61798; S61880
R:Mayer, K.; Wolff, E.; Clevers, H.; Ballhausen, W.G.
Biochim. Biophys. Acta 1263, 169-172, 1995
A:Title: The human high mobility group (HMG)-box transcription factor TCF-1: novel isoform
A:Reference number: S61798; MUID:95367594; PMID:7640309
A:Accession: S61798
A:Molecule type: mRNA
A:Residues: 1-13 <WAS>
A:Cross-references: EMBL:247364
A>Note: DNA was also sequenced
C:Keywords: alternative splicing; DNA binding; transcription factor

Query Match 64.5%; Score 20; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 6.9e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
|
|
Db 6 WDGW 9

RESULT 6

PHI322
IG heavy chain DJ region (clone C344-99) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI322
R:Wasserman, R.; Galli, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
A:Reference number: PHI322; MUID:93094761; PMID:1460419
A:Accession: PHI322
A:Molecule type: DNA
A:Residues: 1-14 <WAS>
C:Keywords: heterotetramer; immunoglobulin

Query Match 64.5%; Score 20; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 7.4e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
|
|
Db 6 WDWY 9

RESULT 7

S23372
T-cell receptor alpha chain J region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999
C:Accession: S23372
R:Pluschke, G.; Ricken, G.; Taube, H.; Kroninger, S.; Melchers, I.; Peter, H.H.; Eich
Eur. J. Immunol. 21, 2749-2754, 1991
A:Title: Biased T cell receptor V(alpha) region repertoire in the synovial fluid of r
A:Reference number: S23364; MUID:92037820; PMID:1657615
A:Accession: S23372
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-13 <PUU>
A:Cross-references: EMBL:X58167
C:Keywords: T-cell receptor

Query Match 54.8%; Score 17; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WXF 6
|
|
Db 11 WTF 13

RESULT 8

B25448
Ig kappa-1 chain, 69 allotype, J-K1.1 segment - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 05-Nov-1999
C:Accession: B25448
R:Akimenko, M.A.; Mariame, B.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 83, 5180-5183, 1986
A:Title: Evolution of the immunoglobulin kappa light chain locus in the rabbit: evide
A:Reference number: A94110; MUID:86259753; PMID:3088570
A:Accession: B25448
A:Molecule type: DNA
A:Residues: 1-13 <AKI>
A:Cross-references: GB:M14067; GB:M14062; GB:M14063; GB:M14064; GB:M14065; GB:M14066;
C:Keywords: heterotetramer; immunoglobulin

Query Match 54.8%; Score 17; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WXF 6

Db | |
1 WAF 3

RESULT 9
B26406
Ig kappa chain J region - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Aug-1996
C:Accession: B26406
R:Sanz, I.; Capra, J.D.
Proc. Natl. Acad. Sci. U.S.A. 84, 1085-1089, 1987
A:Title: V-K and J-K gene segments of A/J Ars-A antibodies: somatic recombination generated
A:Reference number: A26406; MUID:87147197; PMID:3103124
A:Accession: B26406
A:Molecule type: DNA
A:Residues: 1-13 <S>
A:Cross-references: GB:M15519
C:Keywords: heterotetramer; immunoglobulin

Query Match 54.8%; Score 17; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| |
1 WTF 3

Db

RESULT 10
A47630
Ig kappa chain J region J1 - southeastern Australian rat
C:Species: Rattus sordidus villosissimus (southeastern Australian rat)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 05-Nov-1999
C:Accession: A47630
R:Gutman, G.A.; Besta, R.M.; Frank, M.B.; Baverstock, P.R.
Immunogenetics 26, 14-20, 1987
A:Title: Duplication of J-kappa genes within genus Rattus.
A:Reference number: A47630; MUID:87278355; PMID:3111993
A:Accession: A47630
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-13 <GT>
A:Cross-references: GB:M47319; NID:g204788; PIDN:AAA41397.1; PID:g204789
C:Keywords: heterotetramer; immunoglobulin

Query Match 54.8%; Score 17; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| |
1 WTF 3

Db

RESULT 11
T13818
Cytochrome oxidase subunit I - Atlantic hagfish mitochondrial (fragment)
C:Species: mitochondrial Myxine glutinosa (Atlantic hagfish)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13818
R:Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A:Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
A:Reference number: Z17775; MUID:97398704; PMID:9254918
A:Accession: T13818
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-8
A:Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340022
C:Genetics:
A:Genome: mitochondrion
A>Note: COI

C:Keywords: mitochondrion

Query Match 51.6%; Score 16; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| |
6 WFF 8

Db

RESULT 12
T17054
cytochrome-c oxidase (EC 1.9.3.1) chain I - Basiliiscus plumifrons mitochondrion (fragment)
C:Species: mitochondrion Basiliiscus plumifrons
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 22-Oct-1999
C:Accession: T17054
R:Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.
J. Mol. Evol. 44, 660-674, 1997
A:Title: Evolutionary shifts in three major structural features of the mitochondrial g
A:Reference number: Z18674; MUID:97315309; PMID:9189559
A:Accession: T17054
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-10 <MAC>
A:Cross-references: EMBL:U82680; NID:g3603104; PID:g3603107; PIDN:AAC62269.1
C:Genetics:
A:Genome: mitochondrion
A>Note: COI
C:Keywords: mitochondrion; oxidoreductase

Query Match 51.6%; Score 16; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.5e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| |
6 WLF 8

Db

RESULT 13
T13976
cytochrome-c oxidase (EC 1.9.3.1) chain I - Cnemidophorus tigris mitochondrion (fragment)
C:Species: mitochondrion Cnemidophorus tigris
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T13976
R:Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.
Mol. Biol. Evol. 14, 91-104, 1997
A:Title: Two novel gene orders and the role of light-strand replication in rearrangeme
A:Reference number: Z17789; MUID:97153826; PMID:9000757
A:Accession: T13976
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-10 <MAC>
A:Cross-references: EMBL:U71332; NID:gl753236; PID:gl753239; PIDN:AAB48274.1
C:Genetics:
A:Genome: mitochondrion
A>Note: COI
C:Keywords: mitochondrion; oxidoreductase

Query Match 51.6%; Score 16; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.5e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| |
6 WFF 8

Db

RESULT 14
T17057
cytochrome-c oxidase (EC 1.9.3.1) chain I - Crotaphytus collaris mitochondrion (fragment)
C:Species: mitochondrion Crotaphytus collaris

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 22-Oct-1999
C;Accession: T17057
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.
J. Mol. Evol. 44, 660-674, 1997
A;Title: Evolutionary shifts in three major structural features of the mitochondrial gene
A;Reference number: Z18674; MUID:97315309; PMID:9169559
A;Accession: T17057
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: EMBL:U82681; NID:g3603108; PID:g3603111; PIDN:AAC62272.1
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 51.6%; Score 16; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.5e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| |
Db 6 WFF 8

RESULT 15

T12303
cytochrome-c oxidase (EC 1.9.3.1) chain I - Diposaurus dorsalis mitochondrion (fragment)
C;Species: mitochondrion Diposaurus dorsalis
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 22-Oct-1999
C;Accession: T12303
R;Schulte, J.A.; Macey, J.R.; Larson, A.; Papenfuss, T.J.
Mol. Phylogenet. Evol. 10, 367-376, 1998
A;Title: Molecular tests of phylogenetic taxonomies: A general procedure and example using
A;Reference number: Z17488; MUID:95162288; PMID:10051389
A;Accession: T12303
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <SCH>
A;Cross-references: EMBL:AF049857; NID:g4105726; PID:g4105729; PIDN:AAD02514.1
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 51.6%; Score 16; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.5e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| |
Db 6 WFF 8

Search completed: June 10, 2004, 11:07:04
Job time : 11.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 11:01:15 ; Search time 7.66667 Seconds
(without alignments)
40.751 Million cell updates/sec

Title: US-09-912-414-9

Perfect score: 31

Sequence: 1 WXXWF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 801

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	21	67.7	10	1 LABA_JATMU	P13270 jatropha mu
2	19	61.3	9	1 COW_CONVE	P83047 conus ventr
3	14	45.2	9	1 LITR_PHYRO	P08946 phyllomedus
4	14	45.2	10	1 GONI_ALIMI	P37041 alligator m
5	14	45.2	10	1 GON3_ONCKE	P20367 oncorhynch
6	14	45.2	11	1 RANC_RANPI	P08951 rana pipien
7	13	41.9	8	1 RT34_BOVIN	P82929 bos taurus
8	13	41.9	9	1 LITO_LITAU	P08945 litoria aur
9	13	41.9	10	1 HTF_TABAT	P14596 tabanus atr
10	13	41.9	12	1 UR2A_CATCO	P04558 catostomus
11	13	41.9	12	1 UR2B_CATCO	P04559 catostomus
12	13	41.9	12	1 UR2B_CYPCA	P04561 cyprinus ca
13	13	41.9	12	1 UR2_GILMI	P01147 gillichthys
14	13	41.9	12	1 UR2_FOLSP	P81022 polyodon sp
15	13	41.9	12	1 UR2_SCYCA	P35490 scyllorhinu
16	13	41.9	13	1 BOML_PSEGU	P42991 pseudophryn
17	12	38.7	6	1 LOK1_LOCM1	P41491 locusta mig
18	12	38.7	8	1 LCK2_LEUMA	P21141 leucophaea
19	12	38.7	8	1 LCK5_LEUMA	P19987 leucophaea
20	12	38.7	8	1 LCK7_LEUMA	P19989 leucophaea
21	12	38.7	10	1 AEGU_AGRAE	P83465 agrocybe ae
22	12	38.7	10	1 CA12_LITCI	P82086 litoria cit
23	12	38.7	10	1 CAER_LITXA	P86264 litoria xan
24	12	38.7	10	1 GONI_CHEPR	P80677 chelyosoma
25	12	38.7	13	1 YNPF_PHOUL	P41122 phorhabdu
26	12	38.7	15	1 RML2_YEAST	P36522 saccharomyc
27	11	35.5	4	1 OCP3_OCTMI	P58649 octopus min
28	11	35.5	5	1 BPP7_BOTIN	P30425 bothrops in
29	11	35.5	5	1 UF01_MOUSE	P38639 mus musculu
30	11	35.5	6	1 E101_LITRU	P82096 litoria rub
31	11	35.5	7	1 BRHP_CONIM	P58803 conus imper
32	11	35.5	7	1 TPFY_PACDA	P83455 pachymedusa
33	11	35.5	7	1 TV51_LITRU	P82065 litoria rub

```

34      11      35.5      7      1      WWA1_ACHFU      P35919 achatina fu
35      11      35.5      7      1      WWA2_ACHFU      P35920 achatina fu
36      11      35.5      7      1      WWA3_ACHFU      P35921 achatina fu
37      11      35.5      8      1      ACI_THUAL      P18691 thunnus alb
38      11      35.5      8      1      AKHG_GRYBI      P14086 gryllus bim
39      11      35.5      8      1      AKH_LIBAU      P25418 libellula a
40      11      35.5      8      1      AKH_MEML      P25423 melolontha
41      11      35.5      8      1      AKH_TABAT      P14595 tabanus atr
42      11      35.5      8      1      CCKN_MACEU      P58785 conus purpu
43      11      35.5      8      1      COW2_CONFU      P04548 periplaneta
44      11      35.5      8      1      HTF1_PERAM      P04549 periplaneta
45      11      35.5      8      1      HTF2_PERAM

```

ALIGNMENTS

```

RESULT 1
LABA_JATMU
ID LABA_JATMU STANDARD; PRT; 10 AA.
AC P13270;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Labaditin.
OS Jatropha multifida (Physic nut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae; Jatropeae;
OC Jatropha.
OX NCBI_TaxID=3996;
RN [1]
RP SEQUENCE.
RC TISSUE=latex;
RA Kosasi S., van der Sluis W.G., Boelens R., T'Hart L.A., Labadie R.P.;
RT "Labaditin, a novel cyclic decapeptide from the latex of Jatropha
RT multifida L. (Euphorbiaceae). Isolation and sequence determination
RT by means of two-dimensional NMR.";
RL FEBS Lett. 256:91-96(1989).
CC -!- FUNCTION: Labaditin is an active peptide which inhibits the
CC classical pathway of complement activation in vitro. Activity
CC seems to be based on an interaction with C1.
CC -!- PTM: This is a cyclic peptide.
CC -!- MISCCELLANEOUS: Latex of this plant is used in folkloric medicine
CC for treatment of infected wounds, skin infections and scabies.
SQ SEQUENCE 10 AA; 1089 MW; D98AADG362D1B362 CRC64;

Query Match 67.7%; Score 21; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
Db 4 WTVW 7

RESULT 2
COW_CONVE
ID COW_CONVE STANDARD; PRT; 9 AA.
AC P83047;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Contryphan-Vn.
OS Conus ventricosus (Mediterranean cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE, SYNTHESIS, DISULFIDE BONDS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=21547785; PubMed=11688995;

```

RA Massilia G.R., Schinina M.E., Ascenzi P., Polticelli F.;
RT "Contryphan-Vn: a novel peptide from the venom of the Mediterranean
snail *Conus ventricosus*.";
RL Biochem. Biophys. Res. Commun. 288:908-913 (2001).
RN [2]
RP STRUCTURE BY NMR, SYNTHESIS, DISULFIDE BONDS, AND FUNCTION.
RX MEDLINE=22532329; PubMed=12646193;
RA Massilia G.R., Eliseo T., Grolleau F., Lapiet B., Barbier J.,
RA Bournaud R., Molgo J., Cicero D.O., Paci M., Schinina M.E.,
RA Ascenzi P., Polticelli F.;
RT "Contryphan-Vn: a modulator of Ca²⁺-dependent K⁺ channels.";
RL Biochem. Biophys. Res. Commun. 303:238-246 (2003).
CC -!- FUNCTION: Affects both voltage-gated and calcium-dependent
potassium channel activities, with composite and diversified
effects in invertebrate and vertebrate systems.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: The cis isomer is the most abundant and is thus thought to be
the functionally relevant conformer.
CC -!- MASS SPECTROMETRY: MW=1088.6; METHOD=MALDI.
CC -!- SIMILARITY: Belongs to the contryphan family.
DR PDB; INXN; 04-MAR-03.
KW Potassium channel inhibitor; Neurotoxin;
KW Potassium channel inhibitor; D-amino acid; Amidation; 3D-structure.
FT DISULFID 3 9
FT MOD_RES 5 5 D-TRYPTOPHAN.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1091 MW; 8D38676323676EBA CRC64;

Query Match 61.3%; Score 19; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXW 4
DB 5 WKPW 8

RESULT 3
LITR PHYRO STANDARD; PRT; 9 AA.
ID LITR PHYRO STANDARD; PRT; 9 AA.
AC P08946;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rhodei-litorin.
OS Phyllomedusa rohdei (Rhode's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;
OC Phyllomedusinae; Phyllomedusa.
OX NCBI_TaxID=8394;
RN [1]
RP SEQUENCE
RC TISSUE=Skin secretion;
RX MEDLINE=85127560; PubMed=3938283;
RA Barra D., Brspaner G.F., Simmaco M., Bossa F., Melchiorri P.,
RA Brspaner V.;
RT "Rhodei-litorin: a new peptide from the skin of *Phyllomedusa rohdei*.";
RL FEBS Lett. 182:53-56 (1985).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ranatensin
family.
DR PIR; S07241; S07241.
DR InterPro; IPR000874; Bombesin.
DR PFam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
KW Amphibian defense peptide; Bombesin family; Amidation;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1090 MW; 4ECCC1861ADC377 CRC64;

RA Massilia G.R., Schinina M.E., Ascenzi P., Polticelli F.;
RT "Contryphan-Vn: a novel peptide from the venom of the Mediterranean
snail *Conus ventricosus*.";
RL Biochem. Biophys. Res. Commun. 288:908-913 (2001).
RN [2]
RP STRUCTURE BY NMR, SYNTHESIS, DISULFIDE BONDS, AND FUNCTION.
RX MEDLINE=22532329; PubMed=12646193;
RA Massilia G.R., Eliseo T., Grolleau F., Lapiet B., Barbier J.,
RA Bournaud R., Molgo J., Cicero D.O., Paci M., Schinina M.E.,
RA Ascenzi P., Polticelli F.;
RT "Contryphan-Vn: a modulator of Ca²⁺-dependent K⁺ channels.";
RL Biochem. Biophys. Res. Commun. 303:238-246 (2003).
CC -!- FUNCTION: Affects both voltage-gated and calcium-dependent
potassium channel activities, with composite and diversified
effects in invertebrate and vertebrate systems.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: The cis isomer is the most abundant and is thus thought to be
the functionally relevant conformer.
CC -!- MASS SPECTROMETRY: MW=1088.6; METHOD=MALDI.
CC -!- SIMILARITY: Belongs to the contryphan family.
DR PDB; INXN; 04-MAR-03.
KW Potassium channel inhibitor; Neurotoxin;
KW Potassium channel inhibitor; D-amino acid; Amidation; 3D-structure.
FT DISULFID 3 9
FT MOD_RES 5 5 D-TRYPTOPHAN.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1091 MW; 8D38676323676EBA CRC64;

Query Match 45.2%; Score 14; DB 1; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 WXXW 6
DB 3 WATGHP 8

RESULT 4
GONI ALLMI STANDARD; PRT; 10 AA.
ID GONI ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I)
(Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (*Alligator mississippiensis*).";
Regul. Pept. 33:105-116 (1991).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
DR PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GNRH.
DR PFam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 45.2%; Score 14; DB 1; Length 10;
Best Local Similarity 33.3%; Pred. No. 2.4e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
DB 3 WSY 5

RESULT 5
GON3 ONCKE STANDARD; PRT; 10 AA.
ID GON3 ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III) (LH-
RH III) (Luliberin III).
OS *Oncorhynchus keta* (Chum salmon), and
OS *Clupea pallasii* (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; *Oncorhynchus*.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O.keta;
RX MEDLINE=83195140; PubMed=6341199;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;

RT "Characterization of a teleost gonadotropin-releasing hormone.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
 RN [2]

RP SEQUENCE, AND FUNCTION.
 RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolefeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
 Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring.";
 RL Endocrinology 141:505-512(2000).

CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 the secretion of both luteinizing and follicle-stimulating
 hormones.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the GnRH family.

DR PIR; A21114; A21114.

DR InterPro; IPR002012; GnRH.

DR Pfam; PF00446; GnRH; 1.

DR PROSITE; PS00473; GnRH; 1.

KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.

FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT MOD_RES 10 10 AMIDATION.

SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

Query Match 45.2%; Score 14; DB 1; Length 10;

Best Local Similarity 33.3%; Pred. No. 2.4e+03;

Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6

Db 3 WSY 5

RESULT 6

ID RANC_RANPI STANDARD; PRT; 11 AA.

AC P08951;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Ranatensin-C.

OS Rana pipiens (Northern leopard frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.

OX NCBI_TaxID=8404;

RN [1]

RP SEQUENCE.

RC TISSUE=Skin secretion;

RX MEDLINE=84131098; PubMed=6141890;

RA Nakajima T.;

RL Unpublished results, cited by:

RL Ersparner V., Ersparner G.F., Mazzanti G., Endean R.;

RL Comp. Biochem. Physiol. 77C:99-108(1984).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Skin.

CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ranatensin

family.

DR InterPro; IPR000874; Bombesin.

DR Pfam; PF02044; Bombesin; 1.

DR PROSITE; PS00257; BOMBESIN; 1.

KW Amphibian defense peptide; Bombesin family; Amidation.

FT MOD_RES 11 11

SQ SEQUENCE 11 AA; 1304 MW; D6C39885A61ADC366 CRC64;

Query Match

Best Local Similarity 45.2%; Score 14; DB 1; Length 11;

Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 5 WATGHF 10

RESULT 7

RT34_BOVIN

ID RT34_BOVIN STANDARD; PRT; 8 AA.

AC P82929;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Mitochondrial 28S ribosomal protein S34 (S34mt) (MRP-S34) (Fragment).

GN MRPS34.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovioidea;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE.

RC TISSUE=Liver;

RX MEDLINE=21276436; PubMed=11279123;

RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;

RT "The small subunit of the mammalian mitochondrial ribosome:

identification of the full complement of ribosomal proteins present.";

RL J. Biol. Chem. 276:19363-19374(2001).

CC -!- SUBUNIT: Component of the mitochondrial ribosome small subunit

(28S) which comprises a 12S rRNA and about 30 distinct proteins.

CC -!- SUBCELLULAR LOCATION: Mitochondrial.

KW Ribosomal protein; Mitochondrion.

FT NON_TER 1 1

FT NON_TER 8 8

SQ SEQUENCE 8 AA; 935 MW; 9639D1A72058637D CRC64;

Query Match

Best Local Similarity 41.9%; Score 13; DB 1; Length 8;

Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 WXXWKF 6

Db 2 WGILTF 7

RESULT 8

LITO_LITAU

ID LITO_LITAU STANDARD; PRT; 9 AA.

AC P08945;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Litorin.

OS Litoria aurea (Green and golden bell frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaoidea; Hylidae;

OC Pelodyadinae; Litoria.

OX NCBI_TaxID=8371;

RN [1]

RP SEQUENCE.

RC TISSUE=Skin secretion;

RX MEDLINE=75187011; PubMed=1140241;

RA Anastasi A., Ersparner V., Endean R.;

RT "Aminoacid composition and sequence of litorin, a bombesin-like

nonapeptide from the skin of the Australian leptodactylid frog

Litoria aurea.";

RL Experientia 31:510-511(1975).

RN [2]

RP SEQUENCE, AND METHYLATION OF GLN-2.

RC TISSUE=Skin secretion;

RX MEDLINE=78003546; PubMed=908397;

RA Anastasi A., Montecucchi P.C., Angelucci F., Ersparner V., Endean R.;

RT "Glu(OMe)3-litorin, the second bombesin-like peptide occurring in

methanol extracts of the skin of the Australian frog Litoria aurea.";

RL Experientia 33:1289-1289(1977).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Skin.

CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ranatensin

```

CC family.
DR PIR; S07204; S07204.
DR PIR; S07205; S07205.
DR InterPro; IPR000874; Bombesin.
DR Pfam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
KW Amphibian defense peptide; Bombesin family; Amidation; Methylation;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 2 2 DEAMIDATION AND METHYLATION (PARTIAL).
FT MOD_RES 9 9 AMIDATION.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1103 MW; D7CC1B862CDC366 CRC64;

Query Match 41.9%; Score 13; DB 1; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 WXXWXF 6
Db 3 WVGHF 8

RESULT 9
HTF_TABAT STANDARD; PRT; 10 AA.
AC P14596;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypertrehalosemic factor (HOTH) (Dipteran corpora cardiaca factor II) (DCC II).
OS Tabanus atratus (Horse fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha; Tabanidae;
OC Tabanus.
OX NCBI_TaxID=7207;
RN [1]
RP TISSUE=Corpora cardiaca;
RX MEDLINE=90046758; PubMed=2813385;
RA Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT Primary structure of two neurotensin hormones with adipokinetic and
RT hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT flies (Diptera).";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC -!- FUNCTION: Hypertrehalosemic factors are neurotensins that
CC elevate the level of trehalose in the hemolymph (trehalose is the
CC major carbohydrate in the hemolymph of insects).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC PIR; B33995; B33995.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neurotensin; Amidation; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1169 MW; 91603678671A9D1 CRC64;

Query Match 41.9%; Score 13; DB 1; Length 10;
Best Local Similarity 33.3%; Pred. No. 3.6e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
Db 8 WGY 10

RESULT 10
UR2A_CATCO STANDARD; PRT; 12 AA.
AC P04558;
DT 13-AUG-1987 (Rel. 05, Created)
DE Urotensin IIA (U-IIA) (White sucker).
OS Catostomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catostomidae; Catostomus.
OX NCBI_TaxID=7971;
RN [1]
RP MEDLINE=84041959; PubMed=6138758;
RX McMaster D., Lederis K.;
RT Isolation and amino acid sequence of two urotensin II peptides from
RT Catostomus commersoni urophyses.";
RL Peptides 4:367-373(1983).
CC -!- FUNCTION: Urotensin is found in the teleost caudal neurosecretory
CC system. It has a suggested role in osmoregulation and as a
CC corticotropin-releasing factor.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC PIR; JS0424; JS0424.
DR InterPro; IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin_II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11
SQ SEQUENCE 12 AA; 1437 MW; 73961BDBB879CEBB CRC64;

Query Match 41.9%; Score 13; DB 1; Length 12;
Best Local Similarity 33.3%; Pred. No. 4.1e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Urotensin IIA (U-IIA) (White sucker).
OS Catostomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catostomidae; Catostomus.
OX NCBI_TaxID=7971;
RN [1]
RP MEDLINE=84041959; PubMed=6138758;
RX McMaster D., Lederis K.;
RT Isolation and amino acid sequence of two urotensin II peptides from
RT Catostomus commersoni urophyses.";
RL Peptides 4:367-373(1983).
CC -!- FUNCTION: Urotensin is found in the teleost caudal neurosecretory
CC system. It has a suggested role in osmoregulation and as a
CC corticotropin-releasing factor.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC PIR; JS0423; JS0423.
DR InterPro; IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin_II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11
SQ SEQUENCE 12 AA; 1336 MW; 969C76DBB879CEBA CRC64;

Query Match 41.9%; Score 13; DB 1; Length 12;
Best Local Similarity 33.3%; Pred. No. 4.1e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
Db 8 WXY 10

RESULT 11
UR2B_CATCO STANDARD; PRT; 12 AA.
AC P04559;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urotensin IIB (U-IIB) (White sucker).
OS Catostomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catostomidae; Catostomus.
OX NCBI_TaxID=7971;
RN [1]
RP MEDLINE=84041959; PubMed=6138758;
RX McMaster D., Lederis K.;
RT Isolation and amino acid sequence of two urotensin II peptides from
RT Catostomus commersoni urophyses.";
RL Peptides 4:367-373(1983).
CC -!- FUNCTION: Urotensin is found in the teleost caudal neurosecretory
CC system. It has a suggested role in osmoregulation and as a
CC corticotropin-releasing factor.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC PIR; JS0424; JS0424.
DR InterPro; IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin_II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11
SQ SEQUENCE 12 AA; 1437 MW; 73961BDBB879CEBB CRC64;

Query Match 41.9%; Score 13; DB 1; Length 12;
Best Local Similarity 33.3%; Pred. No. 4.1e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 4 WXF 6
| :
Db 8 WKY 10

RESULT 12
UR2B_CYPCA STANDARD; PRT; 12 AA.
AC P04561;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urotensin II-beta (UII-beta) (U-II-beta)
OS Cyprinus carpio (Common carp). Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE.
RA Muneata E., Ohtaki T., Ichikawa T., McMaster D., Lederis K.;
RL (In) Rich D.H., Gross E. (eds.);
RL Proceedings of the 7th American peptide symposium, pp.69-72,
RL Pierce Chemical Co., Rockford IL. (1981).
CC -!- FUNCTION: Urotensin is found in the teleost caudal neurosecretory
CC system. It has a suggested role in osmoregulation and as a
CC corticotropin-releasing factor.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the urotensin 2 family.
DR InterPro; IPR001483; Urotensin II.
DR Pfam; PF02083; Urotensin II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11 G -> S.
FT VARIANT 2 2
SQ SEQUENCE 12 AA; 1407 MW; 73960A9FB879CEBB CRC64;
Query Match 41.9%; Score 13; DB 1; Length 12;
Best Local Similarity 33.3%; Pred. No. 4.1e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| :
Db 8 WKY 10

RESULT 13
UR2_GILMI STANDARD; PRT; 12 AA.
AC P01147;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urotensin II (U-II) (UII).
OS Gillichthys mirabilis (long-jawed mudsucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorphi; Acanthopterygii; Perciformes; Gobioidae;
OC Gobiidae; Gillichthys.
OX NCBI_TaxID=8222;
RN [1]
RP SEQUENCE.
RA MEDLINE=81054904; PubMed=6107911;
RX Pearson D., Shively J.E., Clark B.R., Geschwind I.I., Barkley M.,
RA Nishioka R., Bern H.A.;
RT "Urotensin II: a sonostatin-like peptide in the caudal
RT neurosecretory system of fishes."
RL Proc. Natl. Acad. Sci. U.S.A. 77:5021-5024(1980).
CC -!- FUNCTION: Urotensin is found in the teleost caudal neurosecretory
CC system. It has a suggested role in osmoregulation and as a
CC corticotropin-releasing factor.
CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the urotensin 2 family.
DR PIR; A01409; UOGM2.
DR PIR; S42765; S42765.
DR InterPro; IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11
SQ SEQUENCE 12 AA; 1384 MW; 968BF8982679CEBA CRC64;
Query Match 41.9%; Score 13; DB 1; Length 12;
Best Local Similarity 33.3%; Pred. No. 4.1e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| :
Db 8 WKY 10

RESULT 14
UR2_POLSP STANDARD; PRT; 12 AA.
AC P81022;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urotensin II (U-II) (UII).
OS Polyodon spathula (North American paddlefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Polyodontidae;
OC Polyodon.
OX NCBI_TaxID=7913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spinal cord;
RX MEDLINE=96051494; PubMed=8536944;
RA Waugh D., Youson J., Mims S.D., Sower S., Conlon J.M.;
RT "Urotensin II from the river lamprey (Lampetra fluviatilis), the sea
RT lamprey (Petromyzon marinus), and the paddlefish (Polyodon
RT spathula).";
RL Gen. Comp. Endocrinol. 99:323-332(1995).
CC -!- FUNCTION: HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC CORTICOTROPIN-RELEASING FACTOR. PROBABLY INVOLVED IN SMOOTH
CC MUSCLE STIMULATION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the urotensin 2 family.
DR InterPro; IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11 BY SIMILARITY
SQ SEQUENCE 12 AA; 1410 MW; 7551E9DB879CEBB CRC64;
Query Match 41.9%; Score 13; DB 1; Length 12;
Best Local Similarity 33.3%; Pred. No. 4.1e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| :
Db 8 WKY 10

RESULT 15
UR2_SCVCA STANDARD; PRT; 12 AA.
AC P35490;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urotensin II (U-II) (UII).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidea; Carchariniiformes;

OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RC TISSUE=Spinal cord;
RX MEDLINE=92319231; PubMed=1620290;
RA Conlon J.M., O'Harte F., Smith D.D., Balment R.J., Hazon N.;
RT "Purification and characterization of urotensin II and parvalbumin
from an elasmobranch fish, Scyliorhinus canicula (common dogfish).";
RL Neuroendocrinology 55:230-235(1992).
CC -!- FUNCTION: HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC CORTICOTROPIN-RELEASING FACTOR. PROBABLY INVOLVED IN SMOOTH
CC MUSCLE STIMULATION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC InterPro; IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin_II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11
SQ SEQUENCE 12 AA; 1526 MW; 804729F9D579CEBA CRC64;

Query Match 41.9%; Score 13; DB 1; Length 12;
Best Local Similarity 33.3%; Pred.No. 4.1e+03;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
| :
Db 8 WKY 10

Search completed: June 10, 2004, 11:07:42
Job time : 8.66667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 11:06:26 ; Search time 31.3333 Seconds
(without alignments)
60.418 Million cell updates/sec

Title: US-09-912-414-9
Perfect score: 31
Sequence: 1 WXXWF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 3954

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rviro:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	67.7	9	Q9R5M1	Q9-5ml staphylococ
2	21	67.7	9	Q38366	Q38366 bacterioph
3	20	64.5	9	Q8SHF0	Q8shf0 chamealeo n
4	20	64.5	12	Q77919	Q77919 pseudotroph
5	20	64.5	13	Q16406	Q16406 homo sapien
6	20	64.5	15	Q53580	Q53580 rhodobacter
7	17	54.8	8	Q94VC1	Q94VC1 varanus rud
8	17	54.8	11	Q94V77	Q94V77 heloderma s
9	17	54.8	13	Q9UDC6	Q9udc6 homo sapien
10	17	54.8	14	Q9SAP8	Q9sap8 pisum sativ
11	16	51.6	8	Q94VF6	Q94vf6 varanus job
12	16	51.6	8	Q8WGD7	Q8wgd7 lonis hitra
13	16	51.6	8	Q94V88	Q94v88 varanus tri
14	16	51.6	8	Q9TD02	Q9td02 terranatos
15	16	51.6	8	Q9T4Y2	Q9t4y2 asterina pe
16	16	51.6	8	Q94VJ4	Q94vj4 varanus ben

17	16	51.6	8	8	Q94V91	Q94v91 varanus tim
18	16	51.6	8	8	Q94VE4	Q94ve4 varanus mel
19	16	51.6	8	8	Q94VF9	Q94vf9 varanus ind
20	16	51.6	8	8	Q9T688	Q9t688 gecko gecko
21	16	51.6	9	8	Q94VH4	Q94vh4 varanus gla
22	16	51.6	9	8	Q94VD8	Q94vd8 varanus nil
23	16	51.6	9	8	Q94VI8	Q94vi8 varanus ere
24	16	51.6	9	8	Q94VJ1	Q94vj1 varanus dor
25	16	51.6	9	8	Q8WGE6	Q8wge6 procambarus
26	16	51.6	9	8	Q94VE1	Q94ve1 varanus mer
27	16	51.6	10	2	Q9T335	Q9t335 acinetobact
28	16	51.6	10	8	Q9T8P3	Q9t8p3 liolaemus a
29	16	51.6	10	8	Q9B4W1	Q9b4w1 triturus v
30	16	51.6	10	8	Q9T8K7	Q9t8k7 liolaemus m
31	16	51.6	10	8	Q9T8N1	Q9t8n1 liolaemus p
32	16	51.6	10	8	Q79903	Q79903 oplurus cuv
33	16	51.6	10	8	Q8WDH0	Q8wdh0 anolis limi
34	16	51.6	10	8	Q8W969	Q8w969 anolis orto
35	16	51.6	10	8	Q8WDH8	Q8wdh8 anolis mest
36	16	51.6	10	8	Q79924	Q79924 elgaria pan
37	16	51.6	10	8	Q9T8T6	Q9t8t6 liolaemus m
38	16	51.6	10	8	Q9T8L3	Q9t8l3 liolaemus l
39	16	51.6	10	8	P92616	P92616 aspidoceli
40	16	51.6	10	8	Q9T8G8	Q9t8g8 liolaemus c
41	16	51.6	10	8	Q9B4X0	Q9b4x0 notophthalm
42	16	51.6	10	8	Q8SHI3	Q8shi3 chamealeo c
43	16	51.6	10	8	Q958K9	Q958k9 rana boylli
44	16	51.6	10	8	Q9TFU9	Q9tfu9 teratoscinc
45	16	51.6	10	8	Q9T8X7	Q9t8x7 phymaturus

ALIGNMENTS

RESULT 1
Q9R5M1 ID Q9R5M1 PRELIMINARY; PRT; 9 AA.
AC Q9R5M1; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 66 kDa cell surface adhesin for heparan sulfate (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE.
RX MEDLINE=92176005; PubMed=1541563;
RA Liang O.D., Ascencio F., Fransson L.A., Wadstrom T.; "
RT "Binding of heparan sulfate to Staphylococcus aureus.";
RL Infect. Immun. 60:899-906(1992).
DR PIR: A43848; A43848.
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 990 MW; 2289DDD7337861B3 CRC64;

Query Match 67.7%; Score 21; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
Db 2 WTGW 5

RESULT 2
Q38366 ID Q38366 PRELIMINARY; PRT; 9 AA.
AC Q38366; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE E gene product (fragment).

OS Bacteriophage phi-X174.
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 OX NCBI_TaxID=10847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88118956; PubMed=2963134;
 RA Buckley K.J., Hayashi M.;
 RT "Role of premature translational termination in the regulation of
 expression of the phiX174 lysis gene.";
 RL J. Mol. Biol. 198; 599-607 (1987).
 DR EMBL; X07809; CAA30668.1; -.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1207 MW; C093B37731B36412 CRC64;
 Query Match 67.7%; Score 21; DB 9; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1e+06;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXW 4
 Db 4 WTLW 7
 RESULT 3
 QBSHF0 PRELIMINARY; PRT; 9 AA.
 AC QBSHF0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN Col.
 OS Chamaeleo namaquensis (Namaqua chameleon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Chamaeleo.
 OX NCBI_TaxID=17917;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Townsend T.M., Larson A.L.;
 RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
 Chamaeleonidae (Reptilia, Squamata).";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF448757; AAU90553.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1205 MW; 358CB72733640733 CRC64;
 Query Match 64.5%; Score 20; DB 8; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1e+06;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXW 4
 Db 2 WLRW 5
 RESULT 4
 O77919 PRELIMINARY; PRT; 12 AA.
 AC O77919;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE MHC class II B locus 4 (Fragment).
 OS Pseudotropheus sp. 'pseudotropheus tropheus complex'.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
 OC Cichlidae; Pseudotropheus.
 OX NCBI_TaxID=51796;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=98315113; PubMed=9649539;
 RA Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
 RA Figueroa F., Sultmann H., Klein J.;
 RT "Linkage relationships and haplotype polymorphism among cichlid mhc
 class II B loci.";
 RL Genetics 149; 1527-1537 (1998).
 DR EMBL; AF050032; AAC41371.1; -.
 FT NON_TER 1
 FT NON_TER 12
 SQ SEQUENCE 12 AA; 1529 MW; 6C2ABFACD5A5B734 CRC64;
 Query Match 64.5%; Score 20; DB 7; Length 12;
 Best Local Similarity 50.0%; Pred. No. 1.9e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXW 4
 Db 1 WDFW 4
 RESULT 5
 Q16406 PRELIMINARY; PRT; 13 AA.
 AC Q16406;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE GHRH-R protein (Fragment).
 GN GHRH-R.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Plrimates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96001284; PubMed=7559877;
 RA Hashimoto K., Koga M., Motomura T., Kasayama S., Koubara H.,
 RA Ohnishi T., Arita N., Hayakawa T., Sato B., Kishimoto T.;
 RT "Identification of alternatively spliced messenger ribonucleic acid
 encoding truncated growth hormone-releasing hormone receptor in human
 pituitary adenomas.";
 RL J. Clin. Endocrinol. Metab. 80; 2933-2939 (1995).
 DR EMBL; S79912; AAD14318.1; -.
 FT NON_TER 1
 SQ SEQUENCE 13 AA; 1612 MW; CE19D7D255D66362 CRC64;
 Query Match 64.5%; Score 20; DB 4; Length 13;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXW 4
 Db 7 WGYW 10
 RESULT 6
 Q53580 PRELIMINARY; PRT; 15 AA.
 AC Q53580;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Light-harvesting complex I alpha polypeptide (Fragment).
 GN PUFA.
 OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
 OC Rhodobacteraceae; Rhodobacter.
 OX NCBI_TaxID=1061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92234963; PubMed=1569029;
 RA Richter P., Brand M., Drews G.;

```

RT "Characterization of LHI- and LHI+ Rhodobacter capsulatus pufa
RT mutants.";
RL J. Bacteriol. 174:3030-3041(1992).
DR EMBL; S97552; AAC60406.1; -.
FT NON TER 15
SQ SEQUENCE 15 AA; 2054 MW; 3561FE413591D31A CRC64;

Query Match 64.5%; Score 20; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 2.2e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
Db 8 WKIW 11

RESULT 7
Q94VC1 PRELIMINARY; PRT; 8 AA.
ID Q94VC1 PRELIMINARY; PRT; 8 AA.
AC Q94VC1 PRELIMINARY; PRT; 8 AA.
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus rudicollis (Rough-necked monitor).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169851;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407521; AAL10116.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 8
SQ SEQUENCE 8 AA; 1053 MW; FE2729D5A36411A6 CRC64;

Query Match 54.8%; Score 17; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WXF 6
Db 4 WSF 6

RESULT 8
Q94V77 PRELIMINARY; PRT; 11 AA.
ID Q94V77 PRELIMINARY; PRT; 11 AA.
AC Q94V77 PRELIMINARY; PRT; 11 AA.
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Heloderma suspectum (Gila monster).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;
OC Heloderma.
OX NCBI_TaxID=8554;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407540; AAL10172.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.

```

```

FT NON TER 11
SQ SEQUENCE 11 AA; 1396 MW; 8E3A6DE0D5A36411 CRC64;

Query Match 54.8%; Score 17; DB 8; Length 11;
Best Local Similarity 66.7%; Pred. No. 5.4e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WXF 6
Db 6 WSF 8

RESULT 9
Q9UDC6 PRELIMINARY; PRT; 13 AA.
ID Q9UDC6 PRELIMINARY; PRT; 13 AA.
AC Q9UDC6 PRELIMINARY; PRT; 13 AA.
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE ENDOTHELIUM-derived RELATING factor, nitric oxide synthase
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054573; PubMed=1385404;
RA Janssens S.P., Simouchi A., Quertermous T., Bloch D.B., Bloch K.D.;
RT "Cloning and expression of a cDNA encoding human endothelium-derived
RT relating factor/nitric oxide synthase.";
RL J. Biol. Chem. 267:22694-22694(1992).
FT NON TER 1
FT NON TER 13
SQ SEQUENCE 13 AA; 1390 MW; 3231B6DFEC7EB867 CRC64;

Query Match 54.8%; Score 17; DB 4; Length 13;
Best Local Similarity 66.7%; Pred. No. 6.2e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WXF 6
Db 1 WAF 3

RESULT 10
Q9SAP8 PRELIMINARY; PRT; 14 AA.
ID Q9SAP8 PRELIMINARY; PRT; 14 AA.
AC Q9SAP8 PRELIMINARY; PRT; 14 AA.
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE LHCP11 (14AA) (Fragment).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=var. Alaska;
RA Dobres M.S., Abler M.L., Thompson W.F.;
RT "Sequence of the 3' untranslated region of a pea.";
RL Nucleic Acids Res. 0:0-0(1988).
DR EMBL; X06822; CAA29970.1; -.
FT NON TER 1
FT NON TER 14
SQ SEQUENCE 14 AA; 1537 MW; D55621E9906EA7AD CRC64;

Query Match 54.8%; Score 17; DB 10; Length 14;
Best Local Similarity 66.7%; Pred. No. 6.6e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WXF 6

```

```

Db      | 4 WAF 6
RESULT 11
Q94VF6  PRELIMINARY;      PRT;      8 AA.
AC Q94VF6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus jobiensis (Peach throat monitor).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169843;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407507; AAL10075.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 8 AA; 1144 MW; EFD729DB436411A6 CRC64;

Query Match      51.6%; Score 16; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 WXF 6
Db      | 4 WYF 6
RESULT 12
Q8WGD7  PRELIMINARY;      PRT;      8 AA.
AC Q8WGD7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome oxidase subunit 1 (Fragment).
OS Lomis hirta.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Lomoidea;
OC Lomidae; Lomis.
OX NCBI_TaxID=177234;
RN [1]
RP SEQUENCE FROM N.A.
RA Morrison C.L.; Harvey A.W.; Lavery S.; Tieu K.; Huang Y.;
RA Cunningham C.W.;
RT "Mitochondrial gene rearrangements support a hypothesis of parallel
RT evolution to the crab-like form.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF436035; AAL1611.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 8 AA; 1038 MW; C5B5B9C733640321 CRC64;

Query Match      51.6%; Score 16; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 WXF 6
Db      | 4 WLF 6
RESULT 13
Q94V88  PRELIMINARY;      PRT;      8 AA.
AC Q94V88;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus tristis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=62052;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407533; AAL10151.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 8 AA; 1041 MW; E8B5B9C7336411A6 CRC64;

Query Match      51.6%; Score 16; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 WXF 6
Db      | 4 WLF 6
RESULT 14
Q9TD02  PRELIMINARY;      PRT;      8 AA.
AC Q9TD02;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
OS Terranatos dolichopterus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Terranatos.
OX NCBI_TaxID=61836;
RN [1]
RP SEQUENCE FROM N.A.
RA Hrbek T.; Larson A.;
RT "The evolution of diapause in the killifish family Rivulidae
RT (Atherinomorpha, Cyprinodontiformes): A molecular phylogenetic and
RT biogeographic perspective.";
RL Evolution 53:1200-1216(1999).
DR EMBL; AF92421; AAF03041.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 8 AA; 1084 MW; F0C9D3640DD4056 CRC64;

Query Match      51.6%; Score 16; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 WXF 6
Db      | 6 WFF 8

```

RESULT 15

Q9T4Y2 PRELIMINARY; PRT; 8 AA.
AC Q9T4Y2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE COI gene product (Fragment).
OS Asterina pectinifera (Starfish).
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX NCBI_TaxID=7594;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89354669; PubMed=2766382;
RA Jacobs H.T., Asakawa S., Araki T., Miura K., Smith M.J., Watanabe K.;
RT "Conserved tRNA gene cluster in starfish mitochondrial DNA.";
RL Curr. Genet. 15:193-206(1989).
DR EMBL; X16886; CAA34767.1; --
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 8
SQ SEQUENCE 8 AA; 1114 MW; F0C9D36415B736D6 CRC64;

Query Match 51.6%; Score 16; DB 8; Length 8;
Best Local Similarity 66.7%; Pred.No. 1e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WXF 6
DB 6 WFF 8

Search completed: June 10, 2004, 11:18:46
Job time : 34.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:51:30 ; Search time 46.3333 Seconds
(without alignments)
36.589 Million cell updates/sec

Title: US-09-912-414-11

Perfect score: 38

Sequence: 1 WXXWHF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 447529

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	92.1	6	3 AAB01505	Aab01505 Peptide w
2	35	92.1	6	3 AAB01506	Aab01506 Peptide w
3	34	89.5	6	3 AAB01499	Aab01499 Peptide w
4	34	89.5	6	3 AAB01492	Aab01492 Peptide w
5	29	76.3	6	6 ABR45594	AbR45594 Staphyloc
6	29	76.3	6	6 ABR45593	AbR45593 Staphyloc
7	29	76.3	14	4 AAM00214	Aam00214 Human ang
8	28	73.7	6	2 AAW28912	Aaw28912 Opioid pe
9	28	73.7	6	2 AAR93770	Aar93770 New pepti
10	28	73.7	6	2 AAY23019	Aay23019 Opioid pe
11	28	73.7	6	3 AAB01509	Aab01509 Peptide w
12	28	73.7	6	6 ABR45592	AbR45592 Staphyloc
13	28	73.7	6	6 ABR45591	AbR45591 Staphyloc
14	28	73.7	7	4 AAY01258	Aay01258 US5851813
15	28	73.7	7	4 AAB49729	Aab49729 Peptide S
16	28	73.7	7	6 ABR75385	AbR75385 Biologica
17	28	73.7	7	6 ABR75386	AbR75386 Biologica
18	28	73.7	7	7 ADC27840	Adc27840 Synthetic
19	28	73.7	8	2 AAR60429	Aar60429 Antiprol
20	28	73.7	8	2 AAR60444	Aar60444 Antiprol
21	28	73.7	8	2 AAR83499	Aar83499 Zif268 mu
22	28	73.7	8	2 AAY01261	Aay01261 US5851813
23	28	73.7	8	2 AAW84388	Aaw84388 Finger 3
24	28	73.7	9	6 ABR75262	AbR75262 Biologica
25	28	73.7	9	6 ABR75298	AbR75298 Biologica

ALIGNMENTS

RESULT 1

AAB01505

ID AAB01505 standard; peptide; 6 AA.

XX AAB01505;

DT 08-NOV-2000 (first entry)

DE Peptide which binds to transcription factor E2F-1 DNA binding domain.

XX DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.

OS Synthetic.

XX WO2000044771-A1.

PD 03-AUG-2000.

PF 26-JAN-2000; 2000WO-GB000227.

PR 26-JAN-1999; 99GB-00001710.

XX (PROL-) PROLIFIX LTD.

XX Mueller R, Kontermann RE, Montigiani S;

XX WPI; 2000-532806/48.

PT Peptides binding to the DNA binding domain of transcription factor E2F
and inhibiting cell cycle progression, useful for the treatment of
cancer.

XX Example; Page 26; 42pp; English.

XX Peptides which bind to the DNA binding domain of transcription factor E2F
and inhibit cell cycle progression may be useful as research agents to
investigate the interaction between E2F and DP-1, or the activation of
transcription by E2F-1/DP-1 heterodimers. They may also be used for
inducing apoptosis and/or cell cycle arrest in a cell, particularly for
treatment of cancer or other proliferative disorders such as psoriasis
and restenosis

SQ Sequence 6 AA;

Query Match 92.1%; Score 35; DB 3; Length 6;

Best Local Similarity 66.7%; Pred. No. 1.4e+06;

Aay88160 Fluoresce
Aay88108 Fluoresce
Aab60032 Internali
Abr75359 Biologica
Ada8867 Internali
Ade36943 Interfacci
Abr75387 Biologica
Aaw83884 Peptide s
Aam45777 Hil bindi
Abb90493 Hominidae
Aaw32766 Human pla
Aae12188 Polygluta
Abp46201 Human Bly
Abr47161 Staphyloc
Abr45314 Staphyloc
Abr45313 Staphyloc
Abr47162 Staphyloc
Aae26775 Fibrin bi
Aau93672 Granulocy
Aay88154 Oregon gr

26 73.7 12 3 AAY88160
27 73.7 12 3 AAY88108
28 73.7 12 4 AAB60032
29 73.7 12 6 ABR75359
30 73.7 12 7 ADA8867
31 73.7 13 7 ADE36943
32 73.7 14 6 ABR75387
33 71.1 6 2 AAW83884
34 71.1 7 4 AAM45777
35 71.1 8 5 ABB90493
36 71.1 10 2 AAW32766
37 71.1 11 4 AAE12188
38 71.1 13 5 ABP46201
39 68.4 6 6 ABR47161
40 68.4 6 6 ABR45314
41 68.4 6 6 ABR45313
42 68.4 6 6 ABR47162
43 68.4 9 5 AAE26775
44 68.4 9 5 AAU93672
45 68.4 12 3 AAY88154

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
| | | |
Db 1 WAWWHF 6

RESULT 2
AAB01506
ID AAB01506 standard; peptide; 6 AA.
AC AAB01506;
XX
XX 08-NOV-2000 (first entry)
DT
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
XX DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
XX Synthetic.
OS
XX WO200044771-A1.
PN
XX 03-AUG-2000.
PD
XX 26-JAN-2000; 2000WO-GB000227.
PF
XX 26-JAN-1999; 99GB-00001710.
PR
XX (PROL-) PROLIFIX LTD.
PA
XX Mueller R, Kontermann RE, Montigiani S;
PI
XX WPI; 2000-532806/48.
DR
XX Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
XX Example; Page 26; 42pp; English.
PS
XX Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
XX Sequence 6 AA;

Query Match 92.1%; Score 35; DB 3; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
| | | |
Db 1 WVAWHF 6

RESULT 3
AAB01499
ID AAB01499 standard; peptide; 6 AA.
AC AAB01499;
XX
XX 08-NOV-2000 (first entry)
DT
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
XX DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;

KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 2 /note= "Any amino acid"
FT Misc-difference 3
FT Misc-difference 3 /note= "Any amino acid"
XX
XX WO200044771-A1.
PN
XX 03-AUG-2000.
PD
XX 26-JAN-2000; 2000WO-GB000227.
PF
XX 26-JAN-1999; 99GB-00001710.
PR
XX (PROL-) PROLIFIX LTD.
PA
XX Mueller R, Kontermann RE, Montigiani S;
PI
XX WPI; 2000-532806/48.
DR
XX Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
XX Claim 4; Page 9; 42pp; English.
PS
XX Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
XX Sequence 6 AA;

Query Match 89.5%; Score 34; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WXXWHF 6
| | | |
Db 1 WXXWHF 6

RESULT 4
AAB01492
ID AAB01492 standard; peptide; 6 AA.
XX
AC AAB01492;
XX
XX 08-NOV-2000 (first entry)
DT
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
XX DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
KW restenosis.
XX
XX Synthetic.
OS
XX WO200044771-A1.
PN
XX 03-AUG-2000.
PD
XX 26-JAN-2000; 2000WO-GB000227.
PF
XX 26-JAN-1999; 99GB-00001710.
PR

XX (PROL-) PROLIFIX LTD.
 PA Mueller R, Kontermann RE, Montigiani S;
 PI WPI; 2000-532806/48.
 DR
 XX Peptides binding to the DNA binding domain of transcription factor E2F
 PT and inhibiting cell cycle progression, useful for the treatment of
 PT cancer.
 XX Claim 6; Page 2; 42pp; English.
 PS
 XX Peptides which bind to the DNA binding domain of transcription factor E2F
 CC and inhibit cell cycle progression may be useful as research agents to
 CC investigate the interaction between E2F and DP-1, or the activation of
 CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
 CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
 CC treatment of cancer or other proliferative disorders such as psoriasis
 CC and restenosis
 XX Sequence 6 AA;
 SQ

Query Match 89.5%; Score 34; DB 3; Length 6;
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 | | | |
 Db 1 WVRWHF 6

RESULT 5
 ABR45594
 ID ABR45594 standard; peptide; 6 AA.
 XX
 AC ABR45594;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 DE Staphylococcus aureus CHIPS-related peptide #784.
 XX
 XX CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formulated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX WO2003006048-A1.
 PN 23-JAN-2003.
 XX
 XX 11-JUL-2001; 2001WO-EP008004.
 XX
 XX 11-JUL-2001; 2001WO-EP008004.
 PR
 XX (JARI-) JARI PHARM BV.
 PA
 XX Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX WPI; 2003-256333/25.
 DR
 XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases.
 XX

PS Disclosure; Page 13; 89pp; English.
 XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, disease of the central nervous system, gastrointestinal
 CC diseases, skin diseases, genitourinary diseases, joint diseases,
 CC respiratory diseases and HIV infection
 XX Sequence 6 AA;
 SQ

Query Match 76.3%; Score 29; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 | | | |
 Db 1 WTFWYF 6

RESULT 6
 ABR45593
 ID ABR45593 standard; peptide; 6 AA.
 XX
 AC ABR45593;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 DE Staphylococcus aureus CHIPS-related peptide #783.
 XX
 XX CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formulated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX WO2003006048-A1.
 PN 23-JAN-2003.
 XX
 XX 11-JUL-2001; 2001WO-EP008004.
 XX
 XX 11-JUL-2001; 2001WO-EP008004.
 PR
 XX (JARI-) JARI PHARM BV.
 PA
 XX Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX WPI; 2003-256333/25.
 DR
 XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases.
 XX

PS Disclosure; Page 13; 89pp; English.
 XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164
 CC -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from
 CC Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the C5a-
 CC receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils,
 CC monocytes and endothelial cells or involving acute or chronic
 CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, disease of the central nervous system, gastrointestinal
 CC diseases, skin diseases, genitourinary diseases, joint diseases,
 CC respiratory diseases and HIV infection
 XX Sequence 6 AA;
 SQ

CC inflammation reactions. The diseases or disorders include cardiovascular
 CC diseases, disease of the central nervous system, gastrointestinal
 CC diseases, skin diseases, genitourinary diseases, joint diseases,
 CC respiratory diseases and HIV infection

SQ Sequence 6 AA;

Query Match 76.3%; Score 29; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 1 WSPWFYF 6

RESULT 7
 AAM00214
 ID AAM00214 standard; peptide; 14 AA.

AC AAM00214;

DT 01-OCT-2001 (first entry)

DE Human angiopoietin fragment SEQ ID NO: 754.

XX Human; single nucleotide polymorphism; SNP; paternity test;
 KW forensic test; aberrant protein expression.

XX Homo sapiens.

XX WO200151670-A2.

XX 19-JUL-2001.

XX 05-JAN-2001; 2001WO-US000322.

XX 07-JAN-2000; 2000US-0174962P.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach MD;

XX WPI; 2001-451871/48.

XX N-PSDB; AAR89323.

PT Isolated human polynucleotides containing single nucleotide
 PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
 PT infection and diabetes.

PS Disclosure; Page 321; 475pp; English.

XX The present invention relates to human nucleic acids containing single
 CC nucleotide polymorphisms (SNPs). These can be used in forensic and
 CC paternity tests, and to aid in the treatment of diseases associated with
 CC aberrant protein expression, including cancer, amyloidosis, diabetes,
 CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
 CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,
 CC meningitis, muscular disorders, dementia, neurological diseases, tuberculous
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
 CC autoimmunity. The present sequence is a peptide encoded by a
 CC polymorphism-containing oligonucleotide fragment of the invention

XX Sequence 14 AA;

Query Match 76.3%; Score 29; DB 4; Length 14;
 Best Local Similarity 60.0%; Pred. No. 1.6e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
 | | | |
 Db 7 WYTW 11

RESULT 8
 AAW28912
 ID AAW28912 standard; peptide; 6 AA.
 XX
 AC AAW28912;
 XX
 DT 20-JAN-1998 (first entry)
 XX
 DE Opioid peptide.
 XX
 KW enkephalin; mu-opioid receptor ligand; agonist; antagonist.
 XX
 OS Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "N-acetyl-Arg"
 FT Modified-site 6
 FT /note= "the C-terminal is in amide form"

XX US5641861-A.

XX 24-JUN-1997.

XX 07-JUN-1995; 95US-00487006.

XX 07-JUN-1995; 95US-00487006.

XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.

XX Houghten RA, Dooley CT;

XX WPI; 1997-340994/31.

XX New opioid peptide(s) which bind mu receptors specifically - have agonist
 PT or antagonist activity and are used for study and localisation of mu
 PT receptors and to treat peripheral side effects of morphine etc.

PS Disclosure; Col 8; 92pp; English.

XX The patent discloses the following new peptides, which are opioids which
 CC bind specifically to the mu receptor: Ac-Phe-Arg-Trp-Trp-Tyr-Xaa-NH2 (1);
 CC Ac-Arg-Trp-Ile-Gly-Trp-Xaa-NH2 (2); Trp-Trp-Pro-Lys-His-Xaa-NH2 (3); Trp-
 CC Trp-Pro-Xaa-NH2 (4); Tyr-Pro-Phe-Gly-Phe-Xaa-NH2 (5); D-Ile-D-Met-D-Ser-
 CC D-Trp-D-Trp-(Gly)n-Xaa2-NH2 (6); D-Ile-D-Met-D-Thr-D-Trp-Gly-Xaa2-NH2 (7)
 CC ; Tyr-Al-B2-C3-NH2 (214); Pm and red (Me)x(H)y-Tyr-(NMe)z-Tyr-(Xaa3)z-
 CC NH2 (221); and Trp-Trp-Pro-D4-(His)z-(Xaa)z-NH2 (222); where Xaa = any
 CC natural amino acid; Xaa1 = Lys or Arg; n and z = 0 or 1; Xaa2 = Gly or
 CC the D form of any naturally occurring amino acid; Al = D-norvaline or D-
 CC norleucine; B2 = Gly, Phe or Trp; C3 = Trp or naphthylalanine; x and y =
 CC 0-2, but not over 2 in total; Xaa3 = Phe, D-Phe or benzylamino; D4 = Lys
 CC or Arg; Pm and red indicate permethylation and reduction of all CO in
 CC peptide links to methylene. These new compounds are useful: (i) for in
 CC vitro assay and study of opiate receptor subtypes, particularly mu
 CC receptors in the brain; (ii) for in vivo localisation of receptor
 CC subtypes; and (iii) therapeutically to block the peripheral effects (e.g.
 CC constipation and pruritus) of centrally acting pain killers such as
 CC morphine. They are very selective for the mu opioid receptor, over
 CC binding to the delta and kappa receptor subtypes. The present sequence is
 CC a specific example of a peptide (2)

XX Sequence 6 AA;

Query Match 73.7%; Score 28; DB 2; Length 6;
 Best Local Similarity 60.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
 | | | |
 Db 2 WIGWH 6

RESULT 9
AAR93770
ID AAR93770 standard; protein; 6 AA.
XX
AC AAR93770;
XX
DT 23-SEP-1997 (first entry)
XX
DE New peptide which acts as mu-opioid receptor ligand.
XX
KW mu-receptor; opioid; opiate; agonist; antagonist; diagnosis; analgesic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-acetyl-Arg"
FT Misc-difference 6 /note= "this residue is in C-terminal amide form"
FT
XX
PN WO9640208-A1.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US009321.
XX
PR 07-JUN-1995; 95US-00476438.
XX
PA (TORR-) TORREY PINES INST MOLECULAR STUDIES.
XX
PI Dooley CT, Houghten RA;
XX
DR WPI; 1997-051895/05.
XX
PT New mu opioid receptor binding ligand peptide(s) - useful for in-vitro
PT and in-vivo diagnosis, as analgesics, and for blocking peripheral effects
PT of centrally acting drugs, e.g. morphine.
XX
PS Disclosure; Page 19; 57pp; English.
XX
CC The patent discloses eight new groups of opioid peptides which bind to
CC the mu-receptor to act as agonists or antagonists. The peptides can be
CC used for in-vitro assays to study opiate receptor subtypes (especially
CC the mu type) in brain or other tissue samples; and for in-vivo diagnosis
CC to localise opioid subtypes. The peptides are also useful as drugs to
CC treat pathologies associated with other compounds which interact with the
CC opioid receptor system. Therefore they can be used in medicaments for
CC treating pathologies associated with the mu receptor and as analgesics.
CC They can be used therapeutically to block the peripheral effects of
CC centrally acting pain killers, e.g. to prevent side effects such as
CC constipation and pruritis associated with morphine. The present sequence
CC represents a specific example of one of the new groups of peptides, of
CC formula Ac-Arg-Trp-Ile-Gly-Trp-Xaa-NH2 where Xaa = a naturally occurring
CC amino acid
XX
SQ Sequence 6 AA;
Query Match 73.7%; Score 28; DB 2; Length 6;
Best Local Similarity 60.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWH 5
DB 2 WIGWH 6
RESULT 10
AAY23019
ID AAY23019 standard; peptide; 6 AA.
XX
AC AAY23019;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
XX restenosis.
XX
DT 23-AUG-1999 (first entry)
XX
DE Opioid peptide which inhibits binding of enkephalin.
XX
KW Opioid peptide; ligand binding; opioid receptor;
KW micro-selective opioid peptide; enkephalin; opioid receptor system;
KW blocking; peripheral effect; centrally acting pain killer; morphine.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "acetylated"
FT Modified-site 6 /note= "amidated"
FT
XX
PN US919897-A.
XX
PD 06-JUL-1999.
XX
PF 07-JUN-1995; 95US-00488659.
XX
PR 07-JUN-1995; 95US-00488659.
XX
PA (TORR-) TORREY PINES INST MOLECULAR STUDIES.
XX
PI Dooley CT, Houghten RA;
XX
DR WPI; 1999-394647/33.
XX
PT New opioid peptides useful for blocking the peripheral effects of
PT centrally acting pain killers such as morphine.
XX
PS Example 1; Col 8; 92pp; English.
XX
CC The specification describes opioid peptides, in which each of the N atoms
CC in the peptide backbone between respective amino acids is modified by
CC permethylation, perallylation, perethylation, perbenzylation and
CC pernaphthylation. The peptides inhibit ligand binding to an opioid
CC receptor. Specifically, the peptides inhibit the micro-selective opioid
CC peptide enkephalin. The peptides can be used in vivo diagnostically to
CC localize opioid receptor subtypes. They can be used to treat pathologies
CC associated with other compounds which interact with the opioid receptor
CC system. The peptides are especially useful for blocking the peripheral
CC effects of centrally acting pain killers such as morphine. AAY23005-
CC Y23024 represent opioid peptides of the invention, and are derived from
CC the general sequence given in AAY23004
XX
SQ Sequence 6 AA;
Query Match 73.7%; Score 28; DB 2; Length 6;
Best Local Similarity 60.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWH 5
DB 2 WIGWH 6
RESULT 11
AAB01509
ID AAB01509 standard; peptide; 6 AA.
XX
AC AAB01509;
XX
DT 08-NOV-2000 (first entry)
XX
DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX
KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
XX restenosis.
XX

```

OS Synthetic.
FN WO200044771-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000227.
XX
PR 26-JAN-1999; 99GB-00001710.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Mueller R, Kontermann RE, Montigiani S;
XX
DR WPI; 2000-532806/48.
XX
PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX
PS Example; Page 26; 42pp; English.
XX
CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis
XX
SQ Sequence 6 AA;
Query Match 73.7%; Score 28; DB 3; Length 6;
Best Local Similarity 60.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWH 5
DB 1 WYRWH 5
RESULT 12
ABR45592
ID ABR45592 standard; peptide; 6 AA.
XX
AC ABR45592;
XX
DT 10-JUN-2003 (first entry)
XX
DE Staphylococcus aureus CHIPS-related peptide #782.
XX
KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
KW inflammation; cardiovascular disease; central nervous system disease;
KW gastrointestinal disease; skin disease; genitourinary disease;
KW joint disease; respiratory disease; HIV infection; antiinflammatory;
KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
KW gynecological; immunosuppressive; anti-HIV.
XX
OS Staphylococcus aureus.
OS Synthetic.
XX
FN WO2003006048-A1.
XX
PD 23-JAN-2003.
XX
PF 11-JUL-2001; 2001WO-EP008004.
XX
PR 11-JUL-2001; 2001WO-EP008004.
XX
PA (JARI-) JARI PHARM BV.
XX
PI Van Kessel CPM, Gosselaar-De Haas CJC, Kruijtzer JAW;
XX
DR WPI; 2003-256333/25.
XX
PT Combination of peptides derived from chemotaxis inhibiting protein from
PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
PT prophylaxis and treatment of inflammation, cardiovascular, skin and
PT kidney diseases.
PI Van Strijp JAG;
DR WPI; 2003-256333/25.
XX
PT Combination of peptides derived from chemotaxis inhibiting protein from
PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
PT prophylaxis and treatment of inflammation, cardiovascular, skin and
PT kidney diseases.
PI Van Strijp JAG;
DR WPI; 2003-256333/25.
XX
PT Combination of peptides derived from chemotaxis inhibiting protein from
PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
PT prophylaxis and treatment of inflammation, cardiovascular, skin and
PT kidney diseases.

```

XX PS Disclosure; Page 13; 89pp; English.

XX CC The present invention relates to peptides (ABR44811-ABR47162 and ABR47164 -ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS) from *Staphylococcus aureus*. The peptide fragments are useful in the prophylaxis or treatment of diseases or disorders involving the C5a-receptor (C5aR) and/or formylated peptide receptor (FPR) or neutrophils, monocytes and endothelial cells or involving acute or chronic inflammation reactions. The diseases or disorders include cardiovascular diseases, disease of the central nervous system, gastrointestinal diseases, skin diseases, genitourinary diseases, joint diseases, respiratory diseases and HIV infection

XX SQ Sequence 6 AA;

Query Match 73.7%; Score 28; DB 6; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 Db 1 WFFWYF 6

RESULT 14
 AAY01258
 ID AAY01258 standard; peptide; 7 AA.
 XX AC
 XX AC AAY01258;
 XX 01-JUN-1999 (first entry)
 XX DE US5851813 peptide sequence number 45.

XX KW Antigenic composition; primate; lentivirus; nef gene; vaccine; infection;
 KW AIDS; SIVmac239; deletion; mutant.
 XX OS Simian immunodeficiency virus.
 XX OS Synthetic.

XX PN US5851813-A.
 XX 22-DEC-1998.
 XX 27-JAN-1994; 94US-00188583.
 XX 12-JUL-1990; 90US-00551945.
 XX 09-JUL-1991; 91US-00727494.
 XX (HARD) HARVARD COLLEGE.

XX PI Desrosiers RC;
 XX WPI; 1999-080408/07.
 XX N-PSDB; AAX27657.

XX PT Lentivirus antigenic compositions - containing lentivirus with nef gene deletion.

XX PS Disclosure; Fig 5A-B; 93pp; English.

XX CC The invention relates to an antigenic composition comprising an isolated primate lentivirus whose genome contains an engineered non-revertible null mutation in the nef gene, or an infectious DNA clone in a carrier. The antigenic composition is used in vaccines against infection by the lentivirus, e.g. AIDS

XX SQ Sequence 7 AA;

Query Match 73.7%; Score 28; DB 2; Length 7;
 Best Local Similarity 60.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
 Db 1 WEYWH 5

RESULT 15
 AAB49729
 ID AAB49729 standard; peptide; 7 AA.
 XX AC
 XX AC AAB49729;
 XX 10-APR-2001 (first entry)
 XX DE Peptide SEQ ID 40 which binds to the TADG5 protein.

KW TADG5; human; zinc finger; SH3 domain; cell signalling;
 KW cell cycle control.

OS Unidentified.

XX WO200102432-A1.

XX 11-JAN-2001.

XX 30-JUN-2000; 2000WO-US018304.

XX 01-JUL-1999; 99US-00346510.

XX (UYAR-) UNIV ARKANSAS.

XX O'brien TJ, Wang Y;

XX WPI; 2001-123102/13.

XX Novel SH3 domain-containing TADG5 protein useful for regulating gene replication, as a nutrition supplement, and as a marker for human tissue, or in cell cycle control.

XX Example 6; Page 36; 85pp; English.

XX This invention relates to an SH3 domain-containing protein termed TADG5, and its variants. The invention includes amino acid and polynucleotide sequences for TADG5, and oligonucleotides which bind to either the basic amino acid region and/or the zinc finger motif of the TADG5 protein. The basic amino acid region or zinc finger motif of TADG5 is useful for regulating the expression of the TADG5 gene in a cell. The TADG5 protein is useful as a source of amino acids, as a nutrition supplement, and as a marker for human tissue, or in cell cycle control. TADG5 protein or peptides generated from the protein sequence are useful as antigens for the production of polyclonal and monoclonal antibodies. DNA encoding TADG5 is useful as an antisense vehicle for cell cycle control by shutting down signalling or cell division. The present sequence represents a peptide identified from a phage display peptide library through biopanning with the TADG5 protein

XX Sequence 7 AA;

Query Match 73.7%; Score 28; DB 4; Length 7;
 Best Local Similarity 60.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
 Db 3 WMDWH 7

Search completed: June 10, 2004, 11:06:19
 Job time : 47.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:52:15 ; Search time 11.3333 Seconds
(without alignments)
50.925 Million cell updates/sec

Title: US-09-912-414-11

Perfect score: 38

Sequence: 1 WXXWHF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 2522

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	60.5	9	2 S07241	litorin - Rohde's
2	22	57.9	9	2 S07205	litorin 2-Glu - Au
3	22	57.9	9	2 S07204	litorin I - Austr
4	22	57.9	10	2 F49033	T-cell receptor ga
5	22	57.9	13	2 A60409	bombesin-like pept
6	21	55.3	9	2 A43848	cell surface adhes
7	21	55.3	12	2 PH1308	Ig heavy chain DJ
8	20	52.6	12	2 PH1324	Ig heavy chain DJ
9	20	52.6	13	2 S61798	T-cell-specific tr
10	20	52.6	14	2 PH1322	Ig heavy chain DJ
11	18	47.4	9	2 D57444	neuropeptide Grb-A
12	17	44.7	12	2 A29169	phospholipase A2 (
13	17	44.7	15	2 PA0099	phenotypic variati
14	16	42.1	8	2 T13818	cytochrome oxidase
15	16	42.1	10	2 PQ0177	neuromedin C - lau
16	16	42.1	10	2 A60647	neuromedin C - bov
17	16	42.1	10	2 T13976	cytochrome-c oxida
18	16	42.1	10	2 T13976	cytochrome-c oxida
19	16	42.1	10	2 T12303	cytochrome-c oxida
20	16	42.1	10	2 T14019	cytochrome-c oxida
21	16	42.1	10	2 T17060	cytochrome-c oxida
22	16	42.1	10	2 T14043	cytochrome-c oxida
23	16	42.1	10	2 T14054	cytochrome-c oxida
24	16	42.1	10	2 T17066	cytochrome-c oxida
25	16	42.1	10	2 T17069	cytochrome-c oxida
26	16	42.1	10	2 T12308	cytochrome-c oxida
27	16	42.1	10	2 T17072	cytochrome-c oxida
28	16	42.1	10	2 T12312	cytochrome-c oxida
29	16	42.1	10	2 T12316	cytochrome-c oxida

cytochrome-c oxida
cytochrome-c oxida
bombesin - fire-be
proteochondroitin c
glucan 1,3-beta-gl
Ig heavy chain V r
T-cell receptor al
Ig kappa-1 chain,
Ig kappa chain J r
Ig kappa chain J r
leukocyte elastase
dermorphin (trp-4,
neuropeptide FRFA
cytochrome-c oxida
cytochrome-c oxida

ALIGNMENTS

RESULT 1

S07241

litorin - Rohde's leaf frog

C:Species: Phyllomedusa rohdei (Rohde's leaf frog)

C:Date: 12-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 18-Aug-2000

C:Accession: S07241

R:Barra, D.; Falconieri Erspamer, G.; Simmaco, M.; Bossa, F.; Melchiorri, P.; Erspamer,

FEBS Lett. 182, 53-56, 1985

A:Title: Rohde's litorin: a new peptide from the skin of Phyllomedusa rohdei.

A:Reference number: S07241; MUID:85127560; PMID:3938283

A:Accession: S07241

A:Molecule type: protein

A:Residues: 1-9 <BAF>

C:Superfamily: Gastrin-releasing peptide

C:Keywords: amidated carboxyl end; blocked amino end; neuropeptide; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;9/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 60.5%; Score 23; DB 2; Length 9;

Best Local Similarity 50.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWHF 6

DB 3 WATGHF 8

RESULT 2

S07205

litorin 2-Glu - Australian tree frog (Litoria aurea)

C:Species: Litoria aurea

C:Date: 12-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 18-Aug-2000

C:Accession: S07205

R:Anastasi, A.; Montecucchi, P.; Angelucci, F.; Erspamer, V.; Endean, R.

Experientia 33, 1289, 1977

A:Title: Glu(OMe)(2)-litorin, the second bombesin-like peptide occurring in methanol ex

A:Reference number: S07205; MUID:78003546; PMID:908397

A:Accession: S07205

A:Molecule type: protein

A:Residues: 1-9 <ANA>

C:Superfamily: gastrin-releasing peptide

C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;9/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 57.9%; Score 22; DB 2; Length 9;

Best Local Similarity 50.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWHF 6

DB 3 WATGHF 8

```
RESULT 3
S07204
Litorin I - Australian tree frog (Litoria aurea)
C:Species: Litoria aurea
C>Date: 12-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 18-Aug-2000
C:Accession: S07204
R:Anastasi, A.; Erspamer, V.; Endean, R.
Experientia 31, 510-511, 1975
A:Title: Amino acid composition and sequence of litorin, a bombesin-like nonapeptide from
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
A:Reference number: S07204; MUID:75187011; PMID:1140241
A:Accession: S07204
A:Molecule type: protein
A:Residues: 1-9 <ANA>
C:Superfamily: 1-9 gastrin-releasing peptide
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:9/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 57.9%; Score 22; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWHF 6
| | |
Db 3 WVGHP 8

RESULT 4
F49033
T-cell receptor gamma chain V-D-J region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Dec-1993 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: F49033
R:Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.
Eur. J. Immunol. 21, 2999-3007, 1991
A:Title: Functionally distinct subsets of human gamma/delta T cells.
A:Reference number: A49033; MUID:92083926; PMID:1684157
A:Accession: F49033
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-10 <MOR>
A:Cross-references: GB:S72605; NID:g240700; PIDN:AAB20632.1; PID:g240701
A:Note: sequence extracted from NCBI backbone (NCBIN:72605, NCBIP:72606)
C:Keywords: T-cell receptor

Query Match 57.9%; Score 22; DB 2; Length 10;
Best Local Similarity 40.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
| | |
Db 4 WERY 8

RESULT 5
A60409
bombesin-like peptide L - frog (Pseudophryne guentheri)
C:Species: Pseudophryne guentheri
C>Date: 30-Jan-1993 #sequence_revision 30-Jan-1993 #text_change 18-Aug-2000
C:Accession: A60409
R:Simmaco, M.; Severini, C.; De Biase, D.; Barra, D.; Bossa, F.; Roberts, J.D.; Melchior
Peptides 11, 299-304, 1990
A:Title: Six novel tachykinin- and bombesin-related peptides from the skin of the Austr
A:Reference number: A60409; MUID:90287814; PMID:2356157
A:Accession: A60409
A:Molecule type: protein
A:Residues: 1-13 <SIM>
C:Superfamily: 1-13 <SIM>
C:Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:13/Modified site: amidated carboxyl end (Met) #status experimental
```

```
Query Match 57.9%; Score 22; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWHF 6
| | |
Db 7 WVGHP 12

RESULT 6
A43848
cell surface adhesin for heparan sulfate, 66K - Staphylococcus aureus (fragment)
C:Species: Staphylococcus aureus
C>Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 24-Feb-1995
C:Accession: A43848
R:Liaw, O.D.; Ascencio, F.; Fransson, L.A.; Wadstrom, T.
Infect. Immun. 60, 899-906, 1992
A:Title: Binding of heparan sulfate to Staphylococcus aureus.
A:Reference number: A43848; MUID:92176005; PMID:1541563
A:Accession: A43848
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <LIA>
A:Note: sequence extracted from NCBI backbone (NCBIP:85442)

Query Match 55.3%; Score 21; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXW 4
| | |
Db 2 WTGW 5

RESULT 7
PH1308
Ig heavy chain DJ region (clone C731-94) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1308
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lym
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1308
A:Molecule type: DNA
A:Residues: 1-12 <WAS>
C:Keywords: heterotetramer; immunoglobulin

Query Match 55.3%; Score 21; DB 2; Length 12;
Best Local Similarity 40.0%; Pred. No. 6.2e+02;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
| | |
Db 7 WGWGN 11

RESULT 8
PH1324
Ig heavy chain DJ region (clone C510-100) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1324
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lym
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1324
A:Molecule type: DNA
A:Residues: 1-12 <WAS>
C:Keywords: heterotetramer; immunoglobulin
```

Query Match 52.6%; Score 20; DB 2; Length 12;
 Best Local Similarity 50.0%; Pred. No. 9e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXW 4
 | |
 Db 5 WYIW 8

RESULT 9

S61798
 T-cell-specific transcription factor 1 splice form G - human (fragment)
 N/Alternate names: transcription factor TCF-1g
 C/Species: Homo sapiens (man)
 C/Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 24-Jul-1998
 C/Accession: S61798; S61880
 R/Mayer, K.; Wolffe, E.; Clevers, H.; Ballhausen, W.G.
 Biochim. Biophys. Acta 1263, 169-172, 1995
 A/Title: The human high mobility group (HMG)-box transcription factor TCF-1: novel isoform
 A/Reference number: S61796; MUID:95367594; PMID:7640309
 A/Accession: S61798
 A/Molecule type: mRNA
 A/Residues: 1-13 <MAY>
 A/Cross-references: EMBL:247364
 A/Note: DNA was also sequenced
 C/Keywords: alternative splicing; DNA binding; transcription factor

Query Match 52.6%; Score 20; DB 2; Length 13;
 Best Local Similarity 50.0%; Pred. No. 9.7e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXW 4
 | |
 Db 6 WDGW 9

RESULT 10

PHI322
 Ig heavy chain DJ region (clone C344-99) - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C/Accession: PHI322
 R/Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
 J. Exp. Med. 176, 1577-1581, 1992
 A/Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
 A/Reference number: PHI302; MUID:93094761; PMID:1460419
 A/Accession: PHI322
 A/Molecule type: DNA
 A/Residues: 1-14 <WAS>
 C/Keywords: heterotetramer; immunoglobulin

Query Match 52.6%; Score 20; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXW 4
 | |
 Db 6 WDW 9

RESULT 11

D57444
 neuropeptide Grb-AST B4 - two-spotted cricket
 C/Species: Gryllus bimaculatus (two-spotted cricket)
 C/Date: 26-Jan-1996 #sequence_revision 26-Jan-1996 #text_change 26-Jan-1996
 C/Accession: D57444
 R/Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.
 J. Biol. Chem. 270, 21103-21109, 1995
 A/Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cricket
 A/Reference number: A57444; MUID:95403341; PMID:7673141
 A/Accession: D57444
 A/Status: preliminary

A/Molecule type: protein
 A/Residues: 1-9 <LOR>

Query Match 47.4%; Score 18; DB 2; Length 9;
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
 | |
 Db 2 WERFH 6

RESULT 12

A29169
 phospholipase A2 (EC 3.1.1.4) precursor - sheep (fragment)
 C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C/Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 31-Oct-1997
 C/Accession: A29169
 R/Dutilh, C.E.; Van Doren, P.J.; Verheul, F.E.A.M.; De Haas, G.H.
 Eur. J. Biochem. 53, 91-97, 1975
 A/Title: Isolation and properties of phospholipase A2 from ox and sheep pancreas.
 A/Reference number: A94661
 A/Accession: A29169
 A/Molecule type: protein
 A/Residues: 1-12 <DUT>
 C/Superfamily: phospholipase A2
 C/Keywords: carboxylic ester hydrolase; pyroglutamic acid
 F:/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 44.7%; Score 17; DB 2; Length 12;
 Best Local Similarity 66.7%; Pred. No. 2.8e+03;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WHF 6
 | |
 Db 10 WQF 12

RESULT 13

PA0099
 phenotypic variation protein - fungus (Fusarium sporotrichioides) (fragment)
 C/Species: Fusarium sporotrichioides
 C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
 C/Accession: PA0099
 R/Chow, L.P.; Fukaya, N.; Sugura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
 submitted to JIPID, October 1994
 A/Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichi
 A/Reference number: PA0051
 A/Accession: PA0099
 A/Molecule type: protein
 A/Residues: 1-15 <CHO>

Query Match 44.7%; Score 17; DB 2; Length 15;
 Best Local Similarity 66.7%; Pred. No. 3.3e+03;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WHF 6
 | |
 Db 5 WEF 7

RESULT 14

T13818
 cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)
 C/Species: mitochondrion Myxine glutinosa (Atlantic hagfish)
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C/Accession: T13818
 R/Delarbre, C.; Barriol, V.; Tillier, S.; Janvier, P.; Gachelin, G.
 Mol. Biol. Evol. 14, 807-813, 1997
 A/Title: The main features of the craniate mitochondrial DNA between the ND1 and the CO
 A/Reference number: Z17775; MUID:97398704; PMID:9254918
 A/Accession: T13818
 A/Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-8
 A:Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340022
 C:Genetics:
 A:Genome: mitochondrion
 A:Note: COI
 C:Keywords: mitochondrion

Query Match 42.1%; Score 16; DB 2; Length 8;
 Best Local Similarity 66.7%; Pred. NO. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WHF 6
 |
 Db 6 WFF 8

RESULT 15
 PQ0177
 neuromedin C - laughing frog
 C:Species: Rana ridibunda (laughing frog)
 C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 11-Jan-2000
 C:Accession: PQ0177
 R;Conlon, J.M.; O'Harte, F.; Vaudry, H.
 Biochem. Biophys. Res. Commun. 178, 526-530, 1991
 A:Title: Primary structures of the bombesin-like neuropeptides in frog brain show that
 A:Reference number: PQ0177; MUID:91315477; PMID:1859413
 A:Accession: PQ0177
 A:Molecule type: protein
 A:Residues: 1-10 <CON>
 A:Experimental source: brain
 C:Superfamily: gastrin-releasing peptide
 C:Keywords: amidated carboxyl end
 F;10/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 42.1%; Score 16; DB 2; Length 10;
 Best Local Similarity 40.0%; Pred. NO. 3.4e+03;
 Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWH 5
 |
 Db 4 WAVGH 8

Search completed: June 10, 2004, 11:07:05
 Job time : 12.3333 secs

GenCore version 5.1.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 11:01:15 ; Search time 7.66667 Seconds
(without alignments)
40.751 Million cell updates/sec

Title: US-09-912-414-11

Perfect score: 38

Sequence: 1 WXXWHF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 801

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	23	60.5	9	LITR_PHYRO	P08946 phyllomedusa
2	23	60.5	11	RANC_RANPI	P08951 rana pipien
3	22	57.9	9	LITO_LITAU	P08945 litoria aur
4	22	57.9	13	BOML_PSEGU	P42991 pseudophryn
5	21	55.3	10	LABA_JATMU	P13270 jatropha mu
6	19	50.0	9	COW_CONVE	P83847 conus ventr
7	16	42.1	10	GONZ_GHEPR	P80678 chelyosoma
8	16	42.1	10	GRP_RANRI	P23260 rana ridibu
9	16	42.1	14	ALYT_ALYOB	P08944 alytes obst
10	15	39.5	15	RM12_YEAST	P36522 saccharomyc
11	14	36.8	10	FARP_MYTED	P42560 mytilus edu
12	14	36.8	11	CA22_LITCI	P82088 litoria ctc
13	14	36.8	11	CA42_LITCI	P82092 litoria ctc
14	14	36.8	11	MLG_THETS	P41989 theromyzon
15	14	36.8	13	CXA2_CONGE	P01520 conus geogr
16	14	36.8	13	MLA_ANOCA	P41589 anolis caro
17	14	36.8	13	MLA_CAMDR	P01198 camelus dro
18	14	36.8	15	AH2_PRUSE	P29260 prunus sero
19	14	36.8	15	DCMM_PSECH	P19917 pseudomonas
20	13	34.2	10	APE_CAPGI	P80474 capnocytoph
21	13	34.2	10	GONI_ALIMI	P37041 alligator m
22	13	34.2	10	GON2_CHICK	P37043 gallus gall
23	13	34.2	10	GON3_ONCKE	P20367 oncorhynch
24	13	34.2	12	UR2A_CATCO	P04558 catostomus
25	13	34.2	12	UR2B_CATCO	P04559 catostomus
26	13	34.2	12	UR2B_CXPCA	P04561 cyprinus ca
27	13	34.2	12	UR2_GILMI	P01147 gilllichthys
28	13	34.2	12	UR2_POLSP	P81022 polychdon sp
29	13	34.2	12	UR2_SCVCA	P35490 scyllorhinu
30	13	34.2	15	UC15_MAIZE	P80622 zea mays (m
31	12	31.6	6	LOKI_LOCOMI	P41491 locusta mig
32	12	31.6	8	LCK2_LEUMA	P21141 leucophaea
33	12	31.6	8	LCK5_LEUMA	P19987 leucophaea

34	12	31.6	8	1	LCK7_LEUMA	P19989 leucophaea
35	12	31.6	10	1	AEGL_AGRAE	P83465 agroclybe ae
36	12	31.6	10	1	CA12_LITCI	P82086 litoria ctc
37	12	31.6	10	1	CABR_LITXA	P56264 litoria xan
38	12	31.6	10	1	GON1_CHEPR	P80677 chelyosoma
39	12	31.6	10	1	HTF_TABAT	P14596 tabanus atr
40	12	31.6	11	1	RR2_CONAM	P42341 conopholis
41	12	31.6	15	1	MUP_CAVPO	P83507 cavia porce
42	12	31.6	15	1	RBS_PHYPA	P80657 physcomitre
43	11	28.9	4	1	OCB3_OCTMI	P58649 octopus min
44	11	28.9	5	1	BPF7_BOTIN	P30425 bothrops in
45	11	28.9	5	1	UF01_MOUSE	P38639 mus musculu

ALIGNMENTS

RESULT 1

ID	LITR_PHYRO	STANDARD;	PRT;	9 AA.
AC	P08946;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Rhodei-litorin.			
OS	Phyllomedusa rohdei (Rohde's leaf frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;			
OC	Phyllomedusinae; Phyllomedusa.			
OX	NCBI_TaxID=8394;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Skin secretion;			
RC	MEDLINE=85127560; PubMed=3938283;			
RA	Barra D., Erspamer G.F., Simmaco M., Bossa F., Melchiorri P.,			
RA	Erspamer V.;			
RT	"Rohdei-litorin: a new peptide from the skin of Phyllomedusa rohdei."			
RL	FEBS Lett. 182:53-56(1985).			
CC	- SURCELLULAR LOCATION: Secreted.			
CC	- TISSUE SPECIFICITY: Skin.			
CC	- SIMILARITY: Belongs to the bombesin/neuromedin B/ratanensin			
CC	family.			
CC	PIR; S07241; S07241.			
DR	InterPro: IPR000874; Bombesin.			
DR	Pfam: PF02044; Bombesin; 1.			
DR	PROSITE; PS00257; BOMBESIN; 1.			
DR	Amphibian defense peptide; Bombesin family; Amidation;			
KW	Pyroglutamate carboxylic acid.			
FT	MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.			
FT	MOD_RES 9 9 AMIDATION.			
SQ	SEQUENCE 9 AA; 1090 MW; 4ECCC1E861ADC377 CRC64;			

Query Match 60.5%; Score 23; DB 1; Length 9;

Best Local Similarity 50.0%; Pred. No. 1.4e+05;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 3 WATGHP 8

RESULT 2

ID	RANC_RANPI	STANDARD;	PRT;	11 AA.
AC	P08951;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-NOV-1988 (Rel. 09, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Ranatensin-C.			
OS	Rana pipiens (Northern leopard frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.			
OX	NCBI_TaxID=8404;			

RN SEQUENCE.
 RC TISSUE=Skin secretion;
 RX MEDLINE=84131098; PubMed=6141890;
 RA Nakajima T.;
 RL Unpublished results, cited by:
 RL Erspamer V., Erspamer G.F., Mazzanti G., Eudean R.;
 RL Comp. Biochem. Physiol. 77C:99-108 (1984).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ratanensin family.
 DR InterPro: IPR000874; Bombesin.
 DR Pfam: PF02044; Bombesin; 1.
 DR PROSITE: PS00257; BOMBESIN; 1.
 KW Amphibian defense peptide; Bombesin family; Amidation.
 FT MOD_RES 11 11 AMIDATION.
 SQ SEQUENCE 11 AA; 1304 MW; D6C9885A61ADC366 CRC64;
 Query Match 60.5%; Score 23; DB 1; Length 11;
 Best Local Similarity 50.0%; Pred. No. 88;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 5 WATGHF 10
 RESULT 3
 LITO LITAU
 ID LITO LITAU STANDARD; PRT; 9 AA.
 AC P08945;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Litorin.
 OS Litoria aurea (Green and golden bell frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=8371;
 RN SEQUENCE.
 RC TISSUE=Skin secretion;
 RX MEDLINE=75187011; PubMed=1140241;
 RA Anastasi A., Erspamer V., Eudean R.;
 RT "Amino acid composition and sequence of litorin, a bombesin-like nonapeptide from the skin of the Australian leptodactylid frog Litoria aurea."
 RT Litoria aurea."
 RL Experientia 31:510-511 (1975).
 RN SEQUENCE.
 RP SEQUENCE, AND METHYLATION OF GLN-2.
 RC TISSUE=Skin secretion;
 RX MEDLINE=78003546; PubMed=908397;
 RA Anastasi A., Montecucchi P.C., Angelucci F., Erspamer V., Eudean R.;
 RT "Glu(OMe)3-litorin, the second bombesin-like peptide occurring in methanol extracts of the skin of the Australian frog Litoria aurea."
 RL Experientia 33:1289-1289 (1977).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ratanensin family.
 DR PIR: S07204; S07204.
 DR PIR: S07205; S07205.
 DR InterPro: IPR000874; Bombesin.
 DR Pfam: PF02044; Bombesin; 1.
 DR PROSITE: PS00257; BOMBESIN; 1.
 KW Amphibian defense peptide; Bombesin family; Amidation; Methylation;
 KW Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 2 2 DEAMIDATION AND METHYLATION (PARTIAL).
 FT MOD_RES 9 9 AMIDATION.
 SQ SEQUENCE 9 AA; 1103 MW; D7CCC1E862CDC366 CRC64;

Query Match 57.9%; Score 22; DB 1; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1.4e+05;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 3 WATGHF 8
 RESULT 4
 BOML PSRGU
 ID BOML PSRGU STANDARD; PRT; 13 AA.
 AC P42991;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bombesin-like peptide L (PG-L).
 OS Pseudophryne guentheri (Guenther's toadlet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Myobatrachidae;
 OC Myobatrachinae; Pseudophryne.
 OX NCBI_TaxID=30349;
 RN SEQUENCE.
 RC TISSUE=Skin secretion;
 RX MEDLINE=90287814; PubMed=2356157;
 RA Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
 RA Roberts J.D., Melchiorri P., Erspamer V.;
 RT "Six novel tachykinin- and bombesin-related peptides from the skin of the Australian frog Pseudophryne guntheri."
 RL Peptides 11:299-304 (1990).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ratanensin family.
 DR PIR: A60409; A60409.
 DR InterPro: IPR000874; Bombesin.
 DR Pfam: PF02044; Bombesin; 1.
 DR PROSITE: PS00257; BOMBESIN; 1.
 KW Amphibian defense peptide; Bombesin family; Amidation;
 KW Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 13 13 AMIDATION.
 SQ SEQUENCE 13 AA; 1372 MW; D6DE0D24BD98C366 CRC64;
 Query Match 57.9%; Score 22; DB 1; Length 13;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 7 WATGHF 12
 RESULT 5
 LABA JATMU
 ID LABA JATMU STANDARD; PRT; 10 AA.
 AC P13270;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Labaditin.
 OS Jatropha multifida (Physic nut).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae; Jatrophaeae;
 OC Jatropha.
 OX NCBI_TaxID=3996;
 RN SEQUENCE.
 RP SEQUENCE.
 RC TISSUE=Latex;
 RA Kosasi S., van der Sluis W.G., Boelens R., T'Hart L.A., Labadie R.P.;

RT "Labaditin, a novel cyclic decapeptide from the latex of *Jatropha*
 RT multifida L. (Euphorbiaceae): Isolation and sequence determination
 RL by means of two-dimensional NMR.";
 RL FEMS Lett. 256:91-96(1989).
 CC -!- FUNCTION: Labaditin is an active peptide which inhibits the
 CC classical pathway of complement activation in vitro. Activity
 CC seems to be based on an interaction with C1.
 CC -!- PTM: This is a cyclic peptide.
 CC -!- MISCELLANEOUS: Latex of this plant is used in folkloric medicine
 CC for treatment of infected wounds, skin infections and scabies.
 SQ SEQUENCE 10 AA; 1089 MW; D98AAD6362D1B362 CRC64;

Query Match 55.3%; Score 21; DB 1; Length 10;
 Best Local Similarity 50.0%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 2;

QY 1 WXXW 4
 |
 4 WTVW 7

Db

RESULT 6
 COW CONVE STANDARD; PRT; 9 AA.
 AC P83047;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Contryphan-Vn.
 OS Conus ventricosus (Mediterranean cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=117992;
 RN SEQUENCE, SYNTHESIS, DISULFIDE BONDS, AND MASS SPECTROMETRY.
 RP TISSUE=Venom;
 RC MEDLINE=21547785; PubMed=11688995;
 RX Massilia G.R., Schinina M.E., Ascenzi P., Politicelli F.;
 RA "Contryphan-Vn: a novel peptide from the venom of the Mediterranean
 RT snail *Conus ventricosus*.";
 RL Biochem. Biophys. Res. Commun. 288:908-913(2001).
 RN [2]
 RP STRUCTURE BY NMR, SYNTHESIS, DISULFIDE BONDS, AND FUNCTION.
 RX MEDLINE=2253239; PubMed=12646133;
 RA Massilia G.R., Eliseo I., Grolleau F., Lapied B., Barbier J.,
 RA Bournaud R., Molgo J., Cicero D.O., Paci M., Schinina M.E.,
 RA Ascenzi P., Politicelli F.;
 RT "Contryphan-Vn: a modulator of Ca2+-dependent K+ channels.";
 RL Biochem. Biophys. Res. Commun. 303:238-246(2003).
 CC -!- FUNCTION: Affects both voltage-gated and calcium-dependent
 CC potassium channel activities, with composite and diversified
 CC effects in invertebrate and vertebrate systems.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- PTM: The cis isomer is the most abundant and is thus thought to be
 CC the functionally relevant conformer.
 CC -!- MASS SPECTROMETRY: MW=1088.6; METHOD=MALDI.
 CC -!- SIMILARITY: Belongs to the contryphan family.
 CC PDB; INXN; 04-MAR-03.
 KW Toxin; Ionic channel inhibitor; Neurotoxin;
 KW Potassium channel inhibitor; D-amino acid; Amidation; 3D-structure.
 FT DISULFID 3 9
 FT MOD_RES 5 5 D-TRYPTOPHAN.
 FT MOD_RES 9 9 AMIDATION.
 SQ SEQUENCE 9 AA; 1091 MW; 8D38676323676EBA CRC64;

Query Match 50.0%; Score 19; DB 1; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1.4e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 2;

QY 1 WXXW 4
 |
 4 WTVW 7

Db 5 WKPW 8

RESULT 7
 GON2_CHEPR STANDARD; PRT; 10 AA.
 ID P80678;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GnRH-II)
 DE (Luliberin II).
 OS Chelyosoma productum.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 OC Phlebobranchia; Corellidae; Chelyosoma.
 OX NCBI_TaxID=71177;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96413669; PubMed=8816823;
 RA Powell J.F.F., Reeka-Skinner S.M., Prakash M.O., Fischer W.H.,
 RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
 RT "Two new forms of gonadotropin-releasing hormone in a protochordate
 and the evolutionary implications.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: GnRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
 CC THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
 CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
 CC -!- MASS SPECTROMETRY: MW=1117.52; METHOD=MALDI.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC InterPro; IPR002012; GnRH.
 DR PROSITE; PS00473; GnRH; 1.
 KW Hormone; Amidation; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 6 6 INTERCHAIN.
 FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
 SQ SEQUENCE 10 AA; 1135 MW; 284B38D1EEB735A3 CRC64;

Query Match 42.1%; Score 16; DB 1; Length 10;
 Best Local Similarity 40.0%; Pred. No. 1.2e+03; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 3;

QY 1 WXXW 5
 |
 3 WSLCH 7

Db

RESULT 8
 GRP_RANRI STANDARD; PRT; 10 AA.
 ID P23260;
 AC 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuromedin C.
 OS Rana ridibunda (Laughing frog) (Marsh frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
 OX NCBI_TaxID=8406;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=91315477; PubMed=1859413;
 RA Conlon J.M., O'Harte F., Vaudry H.;
 RT "Primary structures of the bombesin-like neuropeptides in frog brain
 show that bombesin is not the amphibian gastrin-releasing peptide.";
 RL Biochem. Biophys. Res. Commun. 178:526-530(1991).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ratanensin

```

CC PIR; PQ0177; PQ0177.
DR InterPro; IPR000874; Bombesin.
DR Pfam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
DR Bombesin family; Amidation.
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1094 MW; F81FPAE862CDC371 CRC64;

Query Match
Best Local Similarity 42.1%; Score 16; DB 1; Length 10;
Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 4 WAVGH 8

RESULT 9
ALYT_ALYOB
ID ALYT_ALYOB STANDARD; PRT; 14 AA.
AC P08944;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alytesin.
OS Alytes obstetricans (Midwife toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Discoglossidae; Alytes.
OX NCBI_TaxID=8443;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=84131098; PubMed=6141890;
RA Erpamer V., Erpamer G.F., Mazzanti G., Endean R.;
RT "Active peptides in the skins of one hundred amphibian species from
Australia and Papua New Guinea.";
RL Comp. Biochem. Physiol. 77C:99-108(1984).
CC -!- SURCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/zanatinin
family.
DR InterPro; IPR000874; Bombesin.
DR Pfam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
DR Amphibian defense peptide; Bombesin family; Amidation;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1 1
FT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1554 MW; D3C4E4D3AF129666 CRC64;

Query Match
Best Local Similarity 42.1%; Score 16; DB 1; Length 14;
Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 8 WAVGH 12

RESULT 10
RM12 YEAST
ID RM12 YEAST STANDARD; PRT; 15 AA.
AC P3652;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Mitochondrial 60S ribosomal protein L12 (Yml12) (Fragment).
GN MRPL12.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

```

```

RN PIR; PQ0177; PQ0177.
RX MEDLINE=91285106; PubMed=2060626;
RA Grohmann L., Graack H.-R., Kruff V., Choli T., Goldschmidt-Reisin S.,
RA Kitakawa M.;
RT "Extended N-terminal sequencing of proteins of the large ribosomal
subunit from yeast mitochondria.";
RL FEBS Lett. 284:51-56(1991).
DR SGD; I0002687; MRPL12.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1851 MW; 74BCD9FEDDDDB3900 CRC64;

Query Match
Best Local Similarity 39.5%; Score 15; DB 1; Length 15;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 6
DB 1 WXDGYF 6

RESULT 11
FARP MYTED
ID FARP MYTED STANDARD; PRT; 10 AA.
AC P42560;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FMRamide-like neuropeptide ALAGDHFFRF-amide.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=93047883; PubMed=1358534;
RA Walker R.J.;
RT "Neuroactive peptides with an RFamide or Famide carboxyl terminal.";
RL Comp. Biochem. Physiol. 102C:213-222(1992).
CC -!- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
family.
DR PIR; A58365; A58365.
KW Neuropeptide; Amidation.
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1180 MW; C2F80CC9C1EAA87D CRC64;

Query Match
Best Local Similarity 36.8%; Score 14; DB 1; Length 10;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HF 6
DB 6 HF 7

RESULT 12
CA22 LITCI
ID CA22 LITCI STANDARD; PRT; 11 AA.
AC P82088;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Caerulein 2.2/2.2Y4.
OS Litoria citropa (Australian blue mountains tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hyllidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=94770;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=20057701; PubMed=10589099;

```

RA Wabnitz P.A., Bowie J.H., Tyler M.J.;
 RT "Caerulein-like peptides from the skin glands of the Australian blue
 RT contains tree frog *Litoria citropa*. Part 1. Sequence determination
 RT using electrospray mass spectrometry.";
 RL Rapid Commun. Mass Spectrom. 13:2498-2502(1999).
 CC -!- FUNCTION: Hypotensive neuropeptide (Probable).
 CC -!- TISSUE SPECIFICITY: Skin dorsal glands.
 CC -!- PTM: Isoform 2.2Y4 differs from isoform 2.2 in not being
 CC sulfated.
 CC -!- MASS SPECTROMETRY: MW=1389; METHOD=Electrospray.
 CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
 DR INTERPRO: IPR001651; GASTRIN.
 DR PROSITE: PS00259; GASTRIN; FALSE NEG.
 KW Amphibian defense peptide; Hypotensive agent; Amidation; Sulfation;
 KW Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 4 4 SULFATION.
 FT MOD_RES 11 11 AMIDATION.
 FT MOD_RES 11 11 AMIDATION.
 SQ SEQUENCE 11 AA; 1328 MW; 10DAB894EDD861BB CRC64;
 Query Match 36.8%; Score 14; DB 1; Length 11;
 Best Local Similarity 100.0%; Pred. No. 2.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 HF 6
 DB 8 HF 9

RESULT 13
 CA42_LITCI STANDARD; PRT; 11 AA.
 AC P82092;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Caerulein 4.2/4.2Y4.
 OS *Litoria citropa* (Australian blue mountains tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=94770;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Skin secretion;
 RX MEDLINE=20057701; PubMed=10589099;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J.;
 RT "Caerulein-like peptides from the skin glands of the Australian blue
 RT contains tree frog *Litoria citropa*. Part 1. Sequence determination
 RT using electrospray mass spectrometry.";
 RL Rapid Commun. Mass Spectrom. 13:2498-2502(1999).
 CC -!- FUNCTION: Hypotensive neuropeptide (Probable).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin dorsal glands.
 CC -!- PTM: Isoform 4.2Y4 differs from isoform 4.2 in not being
 CC sulfated.
 CC -!- MASS SPECTROMETRY: MW=1404; METHOD=Electrospray.
 CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
 DR INTERPRO: IPR001651; GASTRIN.
 DR PROSITE: PS00259; GASTRIN; FALSE NEG.
 KW Amphibian defense peptide; Hypotensive agent; Amidation; Sulfation;
 KW Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 4 4 SULFATION.
 FT MOD_RES 11 11 AMIDATION.
 FT MOD_RES 11 11 AMIDATION.
 SQ SEQUENCE 11 AA; 1344 MW; 10DAB894F5B861BB CRC64;
 Query Match 36.8%; Score 14; DB 1; Length 11;
 Best Local Similarity 100.0%; Pred. No. 2.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 HF 6

DB 8 HF 9

RESULT 14
 MLG_THETS STANDARD; PRT; 11 AA.
 AC P41989;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Melanotropin gamma (Gamma-melanocyte stimulating hormone) (Gamma-MSH).
 OS Theromyzon tessulatum (leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Rhynchobdellida; Glossiphoniidae; Theromyzon.
 OX NCBI_TaxID=13286;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=94298944; PubMed=8026574;
 RA Salzet M., Watzet C., Bulet P., Malecha J.;
 RT "Isolation and structural characterization of a novel peptide related
 RT to gamma-melanocyte stimulating hormone from the brain of the leech
 RT Theromyzon tessulatum.";
 RL FEBS Lett. 348:102-106(1994).
 CC -!- SIMILARITY: Belongs to the POMC family.
 CC PIR; S45698; S45698.
 KW Hormone; Amidation.
 FT MOD_RES 11 11 AMIDATION.
 SQ SEQUENCE 11 AA; 1486 MW; 2DB8FACE6409C1E8 CRC64;
 Query Match 36.8%; Score 14; DB 1; Length 11;
 Best Local Similarity 100.0%; Pred. No. 2.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 HF 6
 DB 5 HF 6

RESULT 15
 CXA2_CONGE STANDARD; PRT; 13 AA.
 ID CXA2_CONGE
 AC PC1520;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alpha-conotoxin GII.
 OS Conus geographus (Geography cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Caenogastropoda; Sorbeconcha; Hypogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6491;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=81191854; PubMed=7014556;
 RA Gray W.R., Lucue A., Olivera B.M., Barrett J., Cruz L.J.;
 RT "Peptide toxins from Conus geographus venom.";
 RL J. Biol. Chem. 256:4734-4740(1981).
 RN [2]
 RP DISULFIDE BONDS.
 RX MEDLINE=84280842; PubMed=6466616;
 RA Gray W.R., Lucue F.A., Galyean R., Atherton E., Sheppard R.C.,
 RA Stone B.L., Reyes A., Alford J., McIntosh M., Olivera B.M.,
 RA Cruz L.J., Rivier J.;
 RT "Conotoxin GI: disulfide bridges, synthesis, and preparation of
 RT iodinated derivatives.";
 RL Biochemistry 23:2796-2802(1984).
 CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
 CC bind to the nicotinic acetylcholine receptors (nAChR) and thus
 CC inhibit them.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.

CC -!- SIMILARITY: Belongs to the conotoxin A-superfamily. Alpha-type family.

DR PIR; A01783; NTKN2G.

DR HSP; P56973; 1B45.

KW Postsynaptic neurotoxin; Neurotoxin; Toxin;

KW Acetylcholine receptor inhibitor; Amidation.

PT DISULFID 2 7

FT DISULFID 3 13

FT MOD RES 13 13

SQ SEQUENCE 13 AA; 1422 MW; DEER831C39297EBD CRC64;

Query Match 36.8%; Score 14; DB 1; Length 13;

Best Local Similarity 100.0%; Pred. No. 3.4e+03;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HF 6

Db 10 HF 11

Search completed: June 10, 2004, 11:07:42

Job time : 7.66667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 11:06:26 ; Search time 31.3333 Seconds
(without alignments)
60.418 Million cell updates/sec

Title: US-09-912-414-11

Perfect score: 38

Sequence: 1 WXXWHF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 3954

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	57.9	8	13	P79940 xenopus lae
2	21	55.3	9	2	Q9R5M1
3	21	55.3	9	9	Q38366 bacterioph
4	20	52.6	9	8	Q8SHF0 chamealeo n
5	20	52.6	12	7	O77919 pseudotroph
6	20	52.6	13	4	O16406 homo sapien
7	20	52.6	15	2	Q53580 rhodobacter
8	19	50.0	8	8	Q94VF6
9	19	50.0	10	13	Q9PRU9
10	19	50.0	14	8	Q9MT61 allium cepa
11	19	50.0	14	8	Q9MRV4
12	19	50.0	14	8	Q9MRV1
13	19	50.0	14	8	Q9MR18
14	19	50.0	14	8	Q8HGT1 gadus morhu
15	17	44.7	10	8	Q94VD2
16	17	44.7	13	10	Q8LPV3

17	17	44.7	14	6	Q9TQZ1	Q9TQZ1 bos taurus
18	17	44.7	14	11	Q9R1G8	Q9R1G8 rattus norv
19	16	42.1	8	8	Q94VC1	Q94VC1 varanus rud
20	16	42.1	8	8	Q9TD02	Q9TD02 terranatos
21	16	42.1	9	8	Q9T4Y2	Q9T4Y2 asterina pe
22	16	42.1	8	8	Q9T688	Q9T688 gecko gecko
23	16	42.1	10	2	Q47561	Q47561 escherichia
24	16	42.1	10	8	Q9T8K7	Q9T8K7 liolaemus m
25	16	42.1	10	8	Q9T8N1	Q9T8N1 liolaemus p
26	16	42.1	10	8	Q99903	Q99903 oplurus cuv
27	16	42.1	10	8	Q8W969	Q8W969 anolis orto
28	16	42.1	10	8	Q8WDH8	Q8WDH8 anolis mest
29	16	42.1	10	8	Q9T8T6	Q9T8T6 liolaemus m
30	16	42.1	10	8	Q9T8L3	Q9T8L3 liolaemus l
31	16	42.1	10	8	Q92616	Q92616 aspidosceli
32	16	42.1	10	8	Q9T8G8	Q9T8G8 liolaemus c
33	16	42.1	10	8	Q958K9	Q958K9 rana boyllii
34	16	42.1	10	8	Q9T8F9	Q9T8F9 teratoscinc
35	16	42.1	10	8	Q9T8X7	Q9T8X7 phymaturus
36	16	42.1	10	8	Q958L2	Q958L2 rana tempor
37	16	42.1	10	8	Q93885	Q93885 anolis pate
38	16	42.1	10	8	Q9T8Q5	Q9T8Q5 liolaemus l
39	16	42.1	10	8	P92654	P92654 euprepis au
40	16	42.1	10	8	Q9T8L0	Q9T8L0 liolaemus o
41	16	42.1	10	8	Q9T8W8	Q9T8W8 liolaemus b
42	16	42.1	10	8	Q9T8R4	Q9T8R4 liolaemus p
43	16	42.1	10	8	Q9T8M8	Q9T8M8 liolaemus m
44	16	42.1	10	8	Q9T8S1	Q9T8S1 liolaemus l
45	16	42.1	10	8	Q9T8S4	Q9T8S4 liolaemus c

ALIGNMENTS

RESULT 1

P79940 PRELIMINARY; PRT; 8 AA.
AC P79940;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE XMeis1-4 protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97202105; PubMed=9049632;
RA Steelman S., Moskow J.J., Muzynski K., North C., Druck T.,
RA Montgomery J.C., Ruebner K., Daar I.O., Buchberg A.M.;
RT "Identification of a conserved family of Meis1-related homeobox genes.";
RL Genome Res. 7:142-156(1997).
DR EMBL: U68389; AAB19199.1; -.
DR TRANSFAC; T03410; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1187 MW; 278B51F37B11F40B CRC64;

Query Match 57.9%; Score 22; DB 13; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WHF 6
Db 5 WHY 7

RESULT 2

Q9R5M1 PRELIMINARY; PRT; 9 AA.
ID Q9R5M1
AC Q9R5M1;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 66 kDa cell surface adhesin for heparan sulfate (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE.
RX MEDLINE=92176005; PubMed=1541563;
RA Liang O.D., Ascencio F., Fransson L.A., Wadstrom T.;
RT "Binding of heparan sulfate to Staphylococcus aureus";
RL Infect. Immun. 60:899-906(1992).
DR PIR; A43848; A43848.
FT NON TER 1
FT NON TER 9
SQ SEQUENCE 9 AA; 990 MW; 2289DD7337861B3 CRC64;

Query Match 55.3%; Score 21; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06; 2; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
Db 2 WTGW 5

RESULT 3
ID Q38366 PRELIMINARY; PRT; 9 AA.
AC Q38366;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE E gene product (Fragment).
OS Bacteriophage phi-X174.
OC Viruses; ssDNA viruses; Microviridae; Microvirus.
OX NCBI_TaxID=10847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88118956; PubMed=2963134;
RA Buckley K.J., Hayashi M.;
RT "Role of premature translational termination in the regulation of
expression of the phiX174 lysis gene.";
RL J. Mol. Biol. 198:599-607(1987).
DR EMBL; X07809; CAA30668.1; -.
FT NON TER 9
FT NON TER 9
SQ SEQUENCE 9 AA; 1207 MW; C093B37731B36412 CRC64;

Query Match 55.3%; Score 21; DB 9; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06; 2; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
Db 4 WTLW 7

RESULT 4
ID Q8SHF0 PRELIMINARY; PRT; 9 AA.
AC Q8SHF0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
OS COI.
GN Chamaeleo namaquensis (Namaqua chameleon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Chamaeleo.
OX NCBI_TaxID=179917;
RN [1]

RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
Chamaeleonidae (Reptilia, Squamata).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF48757; AAL90553.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 9
FT NON TER 9
SQ SEQUENCE 9 AA; 1205 MW; 358CB72733640733 CRC64;

Query Match 52.6%; Score 20; DB 8; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06; 2; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
Db 2 WLRW 5

RESULT 5
ID O77919 PRELIMINARY; PRT; 12 AA.
AC O77919;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II B locus 4 (Fragment).
OS Pseudotropheus sp. 'pseudotropheus tropheops complex'.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Pseudotropheus.
OX NCBI_TaxID=51796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA Figueroa F., Suitmann H., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid mhc
class II B loci".
RL Genetics 149:1527-1537(1998).
DR EMBL; AF050032; AAC41371.1; -.
FT NON TER 1
FT NON TER 12
FT NON TER 12
SQ SEQUENCE 12 AA; 1529 MW; 6C2ABFACD5A5B734 CRC64;

Query Match 52.6%; Score 20; DB 7; Length 12;
Best Local Similarity 50.0%; Pred. No. 2.6e+03; 0; Mismatches 2; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXW 4
Db 1 WDFW 4

RESULT 6
ID Q16406 PRELIMINARY; PRT; 13 AA.
AC Q16406;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE GHRH-R protein (Fragment).
GN GHRH-R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96001284; PubMed=7559877;
RA Hashimoto K., Koga M., Motomura T., Kasayama S., Kouhara H.,

RA Ohnishi T., Arita N., Hayakawa T., Sato B., Kishimoto T.;
RT "Identification of alternatively spliced messenger ribonucleic acid
RT encoding truncated growth hormone-releasing hormone receptor in human
RL pituitary adenomas.";
RL J. Clin. Endocrinol. Metab. 80:2933-2939 (1995).
DR EMBL: S79912; RAD14318.1; -.
FT NON_TER 1 1
SQ SEQUENCE 13 AA; 1612 MW; CE19D7D25D6362 CRC64;
Query Match 52.6%; Score 20; DB 4; Length 13;
Best Local Similarity 50.0%; Pred. No. 2.8e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXW 4
Db 7 WGYW 10
RESULT 7
Q53580 PRELIMINARY; PRT; 15 AA.
AC Q53580;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Light-harvesting complex I alpha polypeptide (Fragment).
GN PUFA.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92234963; PubMed=1569029;
RA Richter P., Brand M., Drews G.;
RT "Characterization of LHI- and LHI+ Rhodobacter capsulatus pufa
RT mutants.";
RL J. Bacteriol. 174:3030-3041 (1992).
DR EMBL: S97552; AAC60406.1; -.
FT NON_TER 15
SQ SEQUENCE 15 AA; 2054 MW; 3561FE413591D31A CRC64;
Query Match 52.6%; Score 20; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 3.2e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXW 4
Db 8 WKIW 11
RESULT 8
Q94VF6 PRELIMINARY; PRT; 8 AA.
AC Q94VF6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus jobiensis (Peach throat monitor).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguilliforma; Varanidae; Varanus.
OX NCBI_TaxID=169843;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0 (2001).
DR EMBL: AF407507; AAL10075.1; -.
DR GO: GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.

FT NON_TER 8
SQ SEQUENCE 8 AA; 1144 MW; EFD729DB436411A6 CRC64;
Query Match 50.0%; Score 19; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 4 WHF 6
Db 4 WYF 6
RESULT 9
Q9PRU9 PRELIMINARY; PRT; 10 AA.
AC Q9PRU9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Gonadotropin-releasing hormone, SBGNRH-1.
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE.
RX MEDLINE=95083645; PubMed=7991588;
RA Powell J.F., Zohar Y., Elizur A., Park M., Fischer W.H., Craig A.G.,
RA Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
RT brains of one species.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085 (1994).
SQ SEQUENCE 10 AA; 1132 MW; 81568685AB587735 CRC64;
Query Match 50.0%; Score 19; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.2e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 WH 5
Db 8 WH 9
RESULT 10
Q9MT61 PRELIMINARY; PRT; 14 AA.
AC Q9MT61;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PSI 9 kDa protein (Fragment).
GN PSAC.
OS Allium cepa (Onion).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
OC Allium.
OX NCBI_TaxID=4679;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Lopez-Serrano M., del Campo E.M., Sabater B., Martin M.;
RT "Conservation of the start codon by editing in ndnd-encoded
RT transcripts is not restricted to dicotyledonous plants.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ278350; CAB96183.1; -.
DR GO: GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON_TER 1 1
SQ SEQUENCE 14 AA; 1744 MW; 8F14FD03E3B7D911 CRC64;

Query Match 50.0%; Score 19; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.4e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WH 5
 Db 3 WH 4

RESULT 11

Q9MRV4 PRELIMINARY; PRT; 14 AA.
 ID Q9MRV4
 AC Q9MRV4
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE PSI 9 kDa protein (Fragment).
 GN PSAC.
 OS Allium porrum (Leek).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
 OC Allium.
 OC NCBI_TaxID=4681;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Lopez-Serrano M., del Campo E.M., Sabater B., Martin M.;
 RT "Conservation of the start codon by editing in ndhD-encoded
 RT transcripts is not restricted to dicotyledonous plants.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ278352; CAB96185.1; -.
 DR GO; GO:0009507; C:chloroplast; IEA.
 KW Chloroplast.
 FT NON_TER
 SQ SEQUENCE 14 AA; 1744 MW; 8F14FD03E3B7D911 CRC64;

Query Match 50.0%; Score 19; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.4e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WH 5
 Db 3 WH 4

RESULT 12

Q9MRV1 PRELIMINARY; PRT; 14 AA.
 ID Q9MRV1
 AC Q9MRV1
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE PSI 9 kDa protein (Fragment).
 GN PSAC.
 OS Allium sativum (Garlic).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
 OC Allium.
 OC NCBI_TaxID=4682;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Lopez-Serrano M., del Campo E.M., Sabater B., Martin M.;
 RT "Conservation of the start codon by editing in ndhD-encoded
 RT transcripts is not restricted to dicotyledonous plants.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ278351; CAB96187.1; -.
 DR GO; GO:0009507; C:chloroplast; IEA.
 KW Chloroplast.
 FT NON_TER
 SQ SEQUENCE 14 AA; 1744 MW; 8F14FD03E3B7D911 CRC64;

Query Match 50.0%; Score 19; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.4e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WH 5
 Db 3 WH 4

RESULT 13

Q9MRT8 PRELIMINARY; PRT; 14 AA.
 ID Q9MRT8
 AC Q9MRT8
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE PSI 9 kDa protein (Fragment).
 GN PSAC.
 OS Aloe vera (Aloe) (Aloe barbadensis).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Asphodelaceae;
 OC Aloe.
 OC NCBI_TaxID=34199;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Lopez-Serrano M., del Campo E.M., Sabater B., Martin M.;
 RT "Conservation of the start codon by editing in ndhD-encoded
 RT transcripts is not restricted to dicotyledonous plants.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ278353; CAB96192.1; -.
 DR GO; GO:0009507; C:chloroplast; IEA.
 KW Chloroplast.
 FT NON_TER
 SQ SEQUENCE 14 AA; 1744 MW; 8F14FD03E3B7D911 CRC64;

Query Match 50.0%; Score 19; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.4e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WH 5
 Db 3 WH 4

RESULT 14

Q8HGT1 PRELIMINARY; PRT; 14 AA.
 ID Q8HGT1
 AC Q8HGT1
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE ATPase 8 (Fragment).
 GN Gadus morhua (Atlantic cod).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OC NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATPRK3;
 RA Taylor M.I., Fox C., Rico I., Rico C.;
 RT "Species-specific TagMan probes for simultaneous identification of
 RT (Gadus morhua L.), haddock (Melanogrammus aeglefinus L.) and whiting
 RT (Merlangius merlangus L.).";
 RL Mol. Ecol. Notes 2:599-601(2002).
 DR EMBL; AF526615; AAN85062.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER
 SQ SEQUENCE 14 AA; 1744 MW; 8F14FD03E3B7D911 CRC64;

SQ SEQUENCE 14 AA; 1753 MW; D4AF852330085E6D CRC64;
 Query Match 50.0%; Score 19; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.4e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WH 5
 ||
 Db 13 WH 14

RESULT 15
 Q94VD2 PRELIMINARY; PRT; 10 AA.
 AC Q94VD2; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Varanus panoptes panoptes.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
 CX NCBI_TaxID=169849;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ast J.C.;
 RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
 RL Cladistics 17:0-0(2001).
 DR EMBL; AF407516; AAL10102.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1299 MW; 5DEE80D4136411A7 CRC64;

Query Match 44.7%; Score 17; DB 8; Length 10;
 Best Local Similarity 66.7%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WHF 6
 ||
 Db 6 WRF 8

Search completed: June 10, 2004, 11:18:46
 Job time : 31.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:36:04 ; Search time 48 Seconds

(without alignments)

35.318 Million cell updates/sec

Title: US-09-912-414-11

Perfect score: 38

Sequence: 1 WXXWHF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

A_Geneseq 29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	94.7	53	5	Aae21233 Human gen
2	36	94.7	53	5	Abg64892 Human alb
3	35	92.1	6	3	Aab01505 Peptide w
4	35	92.1	6	3	Aab01506 Peptide w
5	35	92.1	36	4	Abb40893 Peptide #
6	35	92.1	36	4	Aam34663 Peptide #
7	35	92.1	36	4	Aam74549 Human bon
8	35	92.1	36	4	Aam61753 Human bra
9	35	92.1	36	4	Abg56337 Human liv
10	35	92.1	65	5	Abp07322 Human ORF
11	35	92.1	113	6	Ada55310 Human pro
12	35	92.1	175	4	Abm03211 Human mus
13	35	92.1	175	4	Aam42371 Human pol
14	35	92.1	175	6	Abu12505 Novel hum
15	35	92.1	256	3	Aag09463 Arabidops
16	35	92.1	256	3	Aag44075 Arabidops
17	35	92.1	256	3	Aag48293 Arabidops
18	35	92.1	258	4	Aag93113 C glutami
19	35	92.1	324	3	Aag50316 Arabidops
20	35	92.1	325	3	Aag48292 Arabidops
21	35	92.1	325	3	Aag44074 Arabidops
22	35	92.1	325	3	Aag09462 Arabidops
23	35	92.1	325	7	Adb95074 A. thalia
24	35	92.1	376	2	Aaw46287 Rhizobium
25	35	92.1	377	2	Aaw46288 Brucella

26	35	92.1	423	3	AAG50315	Arabidops
27	35	92.1	452	3	AAG44023	Arabidops
28	35	92.1	493	3	AAG50314	Arabidops
29	35	92.1	520	3	AAG44022	Arabidops
30	35	92.1	522	3	AAG44021	Arabidops
31	35	92.1	522	6	ABJ25385	Aspergill
32	35	92.1	522	6	ABJ25985	Aspergill
33	35	92.1	6	6	ABJ01499	Peptide w
34	34	89.5	6	3	ABJ01492	Peptide w
35	34	89.5	43	4	AAU14683	Novel bon
36	34	89.5	96	4	AAU14777	Novel bon
37	34	89.5	119	5	ABP06139	Human ORF
38	34	89.5	173	4	AAG83334	P patens
39	34	89.5	173	4	AAG80884	Lipid mod
40	34	89.5	190	6	ADA54839	Human pro
41	34	89.5	306	3	AAG05810	Arabidops
42	34	89.5	306	3	AAG47194	Arabidops
43	34	89.5	338	3	AAG05809	Arabidops
44	34	89.5	338	3	AAG47193	Arabidops
45	34	89.5	402	6	ABJ25408	Aspergill
46	34	89.5	418	2	ABJ25401	Linoleic-
47	34	89.5	423	5	ABB49593	Listeria
48	34	89.5	423	6	ABU32991	Protein e
49	34	89.5	424	4	AAE13431	Soybean m
50	34	89.5	424	6	ABU25617	Protein e
51	34	89.5	427	4	AAU20451	Human sec
52	34	89.5	427	4	AAU21839	Novel hum
53	34	89.5	427	7	ADC46480	Human neo
54	34	89.5	448	3	AAG47192	Arabidops
55	34	89.5	448	3	AAG05808	Arabidops
56	34	89.5	448	5	ABB93179	Herbicida
57	34	89.5	452	7	ADC01384	Enterohae
58	34	89.5	454	5	ABG70026	Larval vi
59	34	89.5	466	4	ABJ95611	Human pro
60	34	89.5	473	5	ABG30731	Human B7-
61	34	89.5	486	7	ABM74221	DNA clone
62	34	89.5	504	6	ABR53206	Protein s
63	34	89.5	506	7	ABD70186	C. neofor
64	34	89.5	511	4	ABB64917	Drosophil
65	34	89.5	511	6	ABR53204	Protein a
66	34	89.5	516	5	ABP73570	Candida a
67	34	89.5	516	5	ABP73631	Candida a
68	34	89.5	529	3	AAJ32200	Human rec
69	34	89.5	591	6	ABJ26008	Aspergill
70	34	89.5	592	6	ABU03036	Moss lipi
71	34	89.5	585	6	ABU39129	Protein e
72	34	89.5	2659	4	ABB68426	Drosophil
73	33	86.8	57	4	AAU66178	Propionib
74	33	86.8	57	6	ABM66059	Propionib
75	33	86.8	57	6	ABM62697	Propionib
76	33	86.8	87	4	AAU56391	Propionib
77	33	86.8	87	6	ABM52910	Propionib
78	33	86.8	140	4	ABJ57913	Drosophil
79	33	86.8	140	4	ABJ71185	Drosophil
80	33	86.8	140	4	ABJ57916	Drosophil
81	33	86.8	142	4	ABJ71187	Drosophil
82	32	84.2	124	5	ABJ78248	Polypepti
83	32	84.2	140	4	ABJ57906	Drosophil
84	32	84.2	210	4	AAU31810	Novel hum
85	32	84.2	214	6	ABU44730	Protein e
86	32	84.2	227	7	ABE62794	Rat Prote
87	32	84.2	227	7	ABE62790	Rat Prote
88	32	84.2	244	5	ABJ93605	Herbicida
89	32	84.2	396	6	AAG98347	Escherich
90	32	84.2	396	6	ABU48344	Protein e
91	32	84.2	396	6	ABU14798	Protein e
92	32	84.2	441	6	ABU38436	Protein e
93	32	84.2	448	4	ABJ50018	Mature ce
94	32	84.2	457	4	ABJ50013	Murine 5H
95	32	84.2	470	4	ABJ50014	Chimeric
96	32	84.2	483	7	ABE61991	Rat Prote
97	32	84.2	483	7	ABE56236	Rat Prote
98	32	84.2	483	7	ABE561995	Rat Prote

99	32	84.2	945	6	ABU17765	Abu17765 Protein e	172	29	76.3	50	4	ABB35142	Abb35142 Peptide #
100	32	84.2	1183	6	ABM70332	Abm70332 Photorhab	173	29	76.3	50	4	ABG50006	Abg50006 Human liv
101	32	84.2	1183	6	ABM70406	Abm70406 Photorhab	174	29	76.3	51	2	AAY38387	Aay38387 Human sec
102	32	84.2	1429	6	ABM69806	Abm69806 Photorhab	175	29	76.3	51	6	ABP73090	Abp73090 Amino aci
103	31	81.6	138	5	AU99675	Aau99675 Mouse nov	176	29	76.3	52	4	AAM20383	Aam20383 Peptide #
104	31	81.6	138	5	AU99666	Aau99666 Human nov	177	29	76.3	52	4	ABB41043	Abb41043 Peptide #
105	31	81.6	147	5	AU99670	Aau99670 Human nov	178	29	76.3	52	4	AAM34818	Aam34818 Peptide #
106	31	81.6	234	4	AAM51001	Aau51001 Propionib	179	29	76.3	52	4	ABB25116	Abb25116 Protein #
107	31	81.6	234	6	ABM47520	Abm47520 Propionib	180	29	76.3	52	4	AAM74702	Aam74702 Human bon
108	31	81.6	383	6	ABU23097	Abu23097 Protein e	181	29	76.3	52	4	AAM61900	Aam61900 Human liv
109	31	81.6	404	3	AAG09658	Aag09658 Arabidops	182	29	76.3	52	4	ABG56486	Abg56486 Human liv
110	31	81.6	408	3	AAG09657	Aag09657 Arabidops	183	29	76.3	52	5	ABG44513	Abg44513 Human pep
111	31	81.6	419	6	ABU42308	Abu42308 Protein e	184	29	76.3	52	6	ABP73093	Abp73093 Amino aci
112	31	81.6	419	6	ABM72005	Abm72005 Scaphyloc	185	29	76.3	52	6	ABP73096	Abp73096 Amino aci
113	31	81.6	420	3	AAG09656	Aag09656 Arabidops	186	29	76.3	52	6	ABP73092	Abp73092 Amino aci
114	31	81.6	485	2	AAM85081	Aaw85081 Esterase	187	29	76.3	54	4	AAO09345	Aao09345 Human pol
115	31	81.6	1286	7	ADC22810	Jdc22810 Human G p	188	29	76.3	55	5	AAB47777	Aab47777 Human thr
116	30	78.9	16	2	AAY14387	Aay14387 Peptide #	189	29	76.3	55	5	ABP05086	Abp05086 Human ORF
117	30	78.9	16	2	AAY15777	Aay15777 Antigenic	190	29	76.3	59	5	ABP07602	Abp07602 Human ORF
118	30	78.9	54	4	AAO13807	Aao13807 Human pol	191	29	76.3	60	5	ABP05584	Abp05584 Human ORF
119	30	78.9	57	3	AAAB3379	Aab3379 Human ORF	192	29	76.3	65	5	ABP32440	Abp32440 Human ORF
120	30	78.9	72	4	AAU59109	Aau59109 Propionib	193	29	76.3	66	6	ABM46003	Abm46003 Propionib
121	30	78.9	72	6	ABM55628	Abm55628 Propionib	194	29	76.3	66	4	AAU49484	Aau49484 Propionib
122	30	78.9	83	4	AAU65122	Aau65122 Propionib	195	29	76.3	69	4	AAU75359	Aau75359 Human col
123	30	78.9	83	6	ABM61641	Abm61641 Propionib	196	29	76.3	69	4	AAU66548	Aau66548 Propionib
124	30	78.9	122	4	AAU40609	Aau40609 Propionib	197	29	76.3	69	6	ABM63067	Abm63067 Propionib
125	30	78.9	122	6	ABM37128	Abm37128 Propionib	198	29	76.3	71	4	AAM16995	Aam16995 Peptide #
126	30	78.9	126	4	AAU40242	Aau40242 Propionib	199	29	76.3	71	4	ABM35984	Abm35984 Peptide #
127	30	78.9	126	4	AAU93345	Aau93345 Human pro	200	29	76.3	71	4	AAM29487	Aam29487 Peptide #
128	30	78.9	126	5	ABM97571	Abm97571 Novel hum	201	29	76.3	71	4	ABM30807	Abm30807 Peptide #
129	30	78.9	126	6	ABM36761	Abm36761 Propionib	202	29	76.3	71	4	ABB21394	Abb21394 Protein #
130	30	78.9	144	4	ABG03034	Abg03034 Novel hum	203	29	76.3	71	4	AAM69164	Aam69164 Human bon
131	30	78.9	147	5	ABM75694	Abm75694 Human DNA	204	29	76.3	71	4	AAM56786	Aam56786 Human bra
132	30	78.9	151	3	ABM53926	Abm53926 Human col	205	29	76.3	71	4	ABG50839	Abg50839 Human liv
133	30	78.9	186	4	ABM17902	Abm17902 Human ner	206	29	76.3	71	4	AAM4702	Aam4702 Peptide #
134	30	78.9	283	4	ABG01564	Abg01564 Novel hum	207	29	76.3	71	5	ABG38761	Abg38761 Human pep
135	30	78.9	382	4	AAU40942	Aau40942 Propionib	208	29	76.3	75	4	AAU45370	Aau45370 Propionib
136	30	78.9	382	6	ABM37461	Abm37461 Propionib	209	29	76.3	75	6	ABM41889	Abm41889 Propionib
137	30	78.9	407	4	ABG20507	Abg20507 Novel hum	210	29	76.3	77	4	AAM80510	Aam80510 Human hae
138	30	78.9	537	4	AAAB79045	Aab79045 Corynebac	211	29	76.3	78	4	ABM03939	Abm03939 Human mus
139	30	78.9	554	4	AAAB80132	Aab80132 Corynebac	212	29	76.3	78	4	ABU031109	Abu031109 Human mus
140	30	78.9	554	4	AAAG91312	Aag91312 C glutami	213	29	76.3	78	6	ABU12403	Abu12403 Novel hum
141	30	78.9	642	4	AAAB79044	Aab79044 Corynebac	214	29	76.3	78	6	ABU13233	Abu13233 Novel hum
142	30	78.9	650	7	ADC31028	Adc31028 Human nov	215	29	76.3	80	4	AAU48729	Aau48729 Propionib
143	30	78.9	689	4	AAAG89921	Aag89921 C glutami	216	29	76.3	80	4	ABM45248	Abm45248 Propionib
144	30	78.9	883	2	AAAB69627	Aab69627 Bacteriop	217	29	76.3	81	5	ABM77724	Abm77724 Amino aci
145	30	78.9	883	2	AAAY01286	Aay01286 Wild-type	218	29	76.3	82	4	ABB15122	Abb15122 Human ner
146	30	78.9	883	2	AAAY01322	Aay01322 Wild-type	219	29	76.3	84	5	ABP06603	Abp06603 Human ORF
147	30	78.9	883	3	AAAB13779	Aab13779 Mutant th	220	29	76.3	97	4	AAM80336	Aam80336 Human hae
148	30	78.9	883	3	AAAB74395	Aab74395 Modified	221	29	76.3	97	4	AAM81756	Aam81756 Human hae
149	30	78.9	883	5	AAO19480	Aao19480 Phage T7	222	29	76.3	100	4	AAU46876	Aau46876 Propionib
150	30	78.9	883	5	AAO19474	Aao19474 Phage T7	223	29	76.3	100	6	ABM43395	Abm43395 Propionib
151	30	78.9	883	5	AAO19474	Aao19474 Phage T7	224	29	76.3	102	4	AAM80410	Aam80410 Human hae
152	30	78.9	883	5	AAO19479	Aao19479 Phage T7	225	29	76.3	107	4	AAO00581	Aao00581 Human pol
153	30	78.9	883	5	AAO19476	Aao19476 Phage T7	226	29	76.3	108	5	ABM90086	Abm90086 Human pol
154	30	78.9	883	5	AAO19477	Aao19477 Phage T7	227	29	76.3	108	6	ADA55602	Ada55602 Human pro
155	30	78.9	883	5	AAO19478	Aao19478 Phage T7	228	29	76.3	109	6	ABU97096	Abu97096 Recombina
156	30	78.9	883	5	AAO19475	Aao19475 Phage T7	229	29	76.3	110	4	AAU50742	Aau50742 Propionib
157	30	78.9	883	5	AAO19481	Aao19481 Phage T7	230	29	76.3	110	6	ABP78465	Abp78465 N. gonorr
158	30	78.9	2004	6	ABP79896	Abp79896 N. gonorr	231	29	76.3	110	6	ABM47261	Abm47261 Propionib
159	30	78.9	2004	6	ABP76777	Abp76777 N. gonorr	232	29	76.3	110	6	ABU97104	Abu97104 Recombina
160	29	76.3	6	6	ABR445594	Abra445594 Scaphyloc	233	29	76.3	112	6	ABB51144	Abb51144 Human sec
161	29	76.3	6	6	ABR445594	Abra445594 Scaphyloc	234	29	76.3	112	6	ABO45401	Abbo45401 Novel hum
162	29	76.3	14	4	AAO00214	Aao00214 Human ang	235	29	76.3	112	7	ABO26881	Abbo26881 Protein a
163	29	76.3	30	4	AAO12524	Aao12524 Human pol	236	29	76.3	115	4	AAM80533	Aam80533 Human hae
164	29	76.3	43	6	AAE30225	Aae30225 Human lp2	237	29	76.3	118	4	AAM87229	Aam87229 Human imm
165	29	76.3	44	4	AAO07688	Aao07688 Human pol	238	29	76.3	118	4	AAM80519	Aam80519 Human hae
166	29	76.3	44	4	ABG55604	Abg55604 Human liv	239	29	76.3	118	4	AAM81803	Aam81803 Human hae
167	29	76.3	44	5	ABG43741	Abg43741 Human pep	240	29	76.3	121	4	AAM80562	Aam80562 Human hae
168	29	76.3	44	7	ADC27591	Adc27591 Human col	241	29	76.3	121	4	AAM81808	Aam81808 Human hae
169	29	76.3	45	4	AAU1936	Aau1936 Novel hum	242	29	76.3	122	4	AAM21750	Aam21750 Peptide #
170	29	76.3	46	4	AAO09235	Aao09235 Human pol	243	29	76.3	122	4	ABB44119	Abb44119 Peptide #
171	29	76.3	50	4	AAM16157	Aam16157 Peptide #	244	29	76.3	122	4	AAM38066	Aam38066 Peptide #

245	29	76.3	122	4	AA080621	Aam80621 Human hae	318	29	76.3	260	5	ABP65326	Abp65326 Bifidobac
246	29	76.3	122	4	AA081810	Aam81810 Human hae	319	29	76.3	265	5	AAU19657	Aau19657 Human nov
247	29	76.3	122	4	AB027006	Ab027006 Protein #	320	29	76.3	285	5	ABP47877	Abp47877 Human pol
248	29	76.3	122	4	AA077846	Aam77846 Human bon	321	29	76.3	285	7	ADC10839	Adc10839 Human ext
249	29	76.3	122	4	AA065141	Aam65141 Human bra	322	29	76.3	267	6	ABM68044	Abm68044 Photorhab
250	29	76.3	122	4	ABG59501	Abg59501 Human liv	323	29	76.3	279	6	AAO30242	Aao30242 Human fib
251	29	76.3	122	4	AA080496	Aam80496 Human hae	324	29	76.3	281	6	ABP79881	Abp79881 N. gonorr
252	29	76.3	123	4	AA081793	Aam81793 Human hae	325	29	76.3	281	6	ABP76835	Abp76835 N. gonorr
253	29	76.3	123	4	AA080728	Aam80728 Human hae	326	29	76.3	282	6	ABR48466	AbR48466 Human Fra
254	29	76.3	123	4	AA081831	Aam81831 Human hae	327	29	76.3	286	6	AAU16730	Aau16730 Protein e
255	29	76.3	124	4	AA081779	Aam81779 Human hae	328	29	76.3	286	2	AAU98520	Aau98520 H. pylori
256	29	76.3	124	4	AA081834	Aam81834 Human hae	329	29	76.3	295	6	ABR43189	AbR43189 Human REM
257	29	76.3	124	4	AA080450	Aam80450 Human hae	330	29	76.3	301	6	AAO30186	Aao30186 Human fib
258	29	76.3	124	4	AA081821	Aam81821 Human hae	331	29	76.3	303	5	ABB89991	Abb89991 Human pol
259	29	76.3	124	4	AA080682	Aam80682 Human hae	332	29	76.3	305	4	ABG09714	Abg09714 Novel hum
260	29	76.3	124	4	AA080764	Aam80764 Human hae	333	29	76.3	309	4	ABG15321	Abg15321 Novel hum
261	29	76.3	127	6	AAU63370	Aau63370 Propionib	334	29	76.3	312	2	AA085760	Aa085760 Human hep
262	29	76.3	127	6	ABM59889	Abm59889 Propionib	335	29	76.3	312	2	AA094317	Aa094317 Hepatocyt
263	29	76.3	127	6	ABM65515	Abm65515 Propionib	336	29	76.3	312	2	AA094316	Aa094316 Hepatocyt
264	29	76.3	132	6	ABP97359	Abp97359 Human ser	337	29	76.3	316	5	AA049496	Aa049496 Lactococ
265	29	76.3	137	4	AAU19813	Aau19813 Human nov	338	29	76.3	328	6	ABU39852	Abu39852 Protein e
266	29	76.3	137	4	ABG03580	Abg03580 Novel hum	339	29	76.3	332	6	ABP78390	Abp78390 N. gonorr
267	29	76.3	137	5	ABP48033	Abp48033 Human pol	340	29	76.3	332	6	ABP76799	Abp76799 N. gonorr
268	29	76.3	137	7	ADC10995	Adc10995 Human pro	341	29	76.3	339	3	ABR43637	AbR43637 Human can
269	29	76.3	139	3	AA040770	Aa040770 Human ORF	342	29	76.3	348	1	AA094682	Aa094682 Humano aci
270	29	76.3	139	4	AB015141	Ab015141 Human ner	343	29	76.3	358	4	AA090341	Aa090341 C glutami
271	29	76.3	139	5	ABP01804	Abp01804 Human ORF	344	29	76.3	361	4	AB066207	Ab066207 Drosophil
272	29	76.3	139	6	AA054690	Aa054690 Human pro	345	29	76.3	365	6	ABP57738	Abp57738 A. thalia
273	29	76.3	147	4	AA090959	Aa090959 Human imm	346	29	76.3	375	6	ABU40624	Abu40624 Protein e
274	29	76.3	148	5	AB053924	Ab053924 Lactococ	347	29	76.3	379	6	ABM67684	Abm67684 Photorhab
275	29	76.3	154	4	AB052719	Ab052719 Escherich	348	29	76.3	381	6	AA067682	Aa067682 Photorhab
276	29	76.3	154	7	AB001212	Ab001212 Enterobac	349	29	76.3	382	6	AA030435	Aa030435 Human hyp
277	29	76.3	155	4	AAU29414	Aau29414 Human G P	350	29	76.3	389	1	AA080300	Aa080300 Pectin es
278	29	76.3	155	5	ABG60702	Abg60702 Novel G P	351	29	76.3	389	2	AA032108	Aa032108 Pectin es
279	29	76.3	162	4	AA090349	Aa090349 Human imm	352	29	76.3	395	6	ABM67014	Abm67014 Photorhab
280	29	76.3	174	7	ADC94576	Adc94576 E. faeciu	353	29	76.3	395	6	ABU28011	Abu28011 Protein e
281	29	76.3	175	4	AA079750	Aa079750 Coryneb	354	29	76.3	395	6	ABU31636	Abu31636 Protein e
282	29	76.3	175	4	AB080015	Ab080015 Coryneb	355	29	76.3	397	6	ABP57732	Abp57732 A. thalia
283	29	76.3	175	4	AB079994	Ab079994 Coryneb	356	29	76.3	399	6	ABU49865	Abu49865 Protein e
284	29	76.3	175	4	AAU71918	Aau71918 C. glutam	357	29	76.3	401	4	AA080014	Aa080014 Coryneb
285	29	76.3	177	7	ADC97284	Adc97284 E. faeciu	358	29	76.3	401	4	AA079993	Ab079993 Coryneb
286	29	76.3	178	4	AAU44711	Aau44711 Propionib	359	29	76.3	401	4	AA079749	Aa079749 Coryneb
287	29	76.3	178	6	ABM41230	Abm41230 Propionib	360	29	76.3	401	4	AAU71917	Aau71917 C. glutam
288	29	76.3	183	4	AAU19348	Aau19348 Human G P	361	29	76.3	401	4	AA092077	Aa092077 C. glutam
289	29	76.3	193	6	ABM69622	Abm69622 Photorhab	362	29	76.3	402	4	ABG05854	Abg05854 Novel hum
290	29	76.3	196	3	AA028193	Aa028193 Arabidops	363	29	76.3	416	7	AD095216	Ad095216 Porphyrom
291	29	76.3	204	6	ABU25570	Abu25570 Protein e	364	29	76.3	416	7	AD095214	Ad095214 Porphyrom
292	29	76.3	205	6	ABM72206	Abm72206 Staphyloc	365	29	76.3	430	4	ABM61244	Abm61244 Drosophil
293	29	76.3	210	4	AAU19818	Aau19818 Human nov	366	29	76.3	430	4	AA079749	Aa079749 Human pro
294	29	76.3	210	5	ABP48038	Abp48038 Human pol	367	29	76.3	474	4	AA079745	Aa079745 A. thalia
295	29	76.3	210	7	ADC11000	Adc11000 Human pro	368	29	76.3	476	6	ABP57745	Abp57745 Enterococ
296	29	76.3	230	3	AA028192	Aa028192 Arabidops	369	29	76.3	504	2	AA000225	Aa000225 Enterococ
297	29	76.3	234	3	AA011120	Aa011120 Arabidops	370	29	76.3	504	5	ABP43444	Abp43444 E. faecal
298	29	76.3	238	3	AB053292	Ab053292 Human col	371	29	76.3	504	6	ABU88472	Abu88472 E. faecal
299	29	76.3	238	7	ADC27598	Adc27598 Human aci	372	29	76.3	504	6	ABU13723	Abu13723 Enterococ
300	29	76.3	241	7	AD070784	Ad070784 Novel pro	373	29	76.3	511	6	ABU03088	Abu03088 Alpha amy
301	29	76.3	245	4	AB063645	Ab063645 Drosophil	374	29	76.3	527	4	AB062569	Ab062569 Drosophil
302	29	76.3	245	3	AA053982	Aa053982 Arabidops	375	29	76.3	527	7	AB070180	Ab070180 C. neofor
303	29	76.3	245	3	AA036996	Aa036996 Arabidops	376	29	76.3	552	2	AA000224	Aa000224 Enterococ
304	29	76.3	247	3	AA0783081	Aa0783081 F-box pro	377	29	76.3	552	5	ABP43443	Abp43443 E. faecal
305	29	76.3	247	3	AA011119	Aa011119 Arabidops	378	29	76.3	552	6	ABU88471	Abu88471 E. faecal
306	29	76.3	247	5	AA022467	Aa022467 Human F-b	379	29	76.3	552	6	ABU13722	Abu13722 Enterococ
307	29	76.3	250	6	ABU97121	Abu97121 Recombina	380	29	76.3	553	4	ABG09100	Abg09100 Novel hum
308	29	76.3	250	6	ABU97147	Abu97147 Recombina	381	29	76.3	569	4	ABG09617	Abg09617 Novel hum
309	29	76.3	251	4	AAU19650	Aau19650 Human nov	382	29	76.3	612	4	ABG16883	Abg16883 Novel hum
310	29	76.3	251	5	ABP47870	Abp47870 Human pol	383	29	76.3	615	6	ABU17205	Abu17205 Protein e
311	29	76.3	251	7	ADC10832	Adc10832 Human ext	384	29	76.3	615	6	ADA33187	Ada33187 Acinetoba
312	29	76.3	252	3	AA011118	Aa011118 Arabidops	385	29	76.3	620	4	ABG05847	Abg05847 Novel hum
313	29	76.3	257	3	AA028191	Aa028191 Arabidops	386	29	76.3	620	4	ABG06486	Abg06486 Novel hum
314	29	76.3	258	6	AA030319	Aa030319 Human lrp2	387	29	76.3	620	4	ABG03595	Abg03595 Novel hum
315	29	76.3	259	3	AA053881	Aa053881 Arabidops	388	29	76.3	620	4	ABG15323	Abg15323 Novel hum
316	29	76.3	259	3	AA036995	Aa036995 Arabidops	389	29	76.3	621	4	AA01891	Aa01891 Arabidops
317	29	76.3	259	5	AB092601	Ab092601 Herbicida	390	29	76.3	621	7	ADD55766	Add55766 Thalecres

391	29	76.3	621	7	ADP31565	Plant yie	Ade31565	29	76.3	1953	7	ADP16058	G-coupled
392	29	76.3	622	6	ABU49786	Protein e	Abu49786	29	76.3	2466	6	ABM70225	Photorhab
393	29	76.3	633	6	ADA35232	Acinetoba	Ada35232	29	76.3	2499	6	ABM70221	Photorhab
394	29	76.3	634	7	ADP16056	G-coupled	Adel16056	29	76.3	2516	2	AAW17899	Photorhab
395	29	76.3	641	4	ABBE1933	Drosophil	Abbe1933	29	76.3	2516	2	AAW56572	Toxin Tcd
396	29	76.3	640	2	ABR70013	Tick carb	Abbr70013	29	76.3	2516	4	AAW72609	Photorhab
397	29	76.3	660	4	ABG00276	Novel hum	Abg00276	29	76.3	2516	5	ABG32651	P. lumine
398	29	76.3	663	4	AAU51590	Propionib	Aau51590	29	76.3	2517	4	AAW72611	Modified
399	29	76.3	663	6	ABM48109	Propionib	Abm48109	29	76.3	2522	2	AAW33729	Photorhab
400	29	76.3	665	6	ABG20291	Novel hum	Abg20291	29	76.3	2526	6	ABM70229	Photorhab
401	29	76.3	668	5	ABW79514	Shc bindi	Abw79514	29	76.3	2537	4	AAW72614	TcdA toxi
402	29	76.3	668	6	ABU08529	Mouse Shc	Abu08529	29	76.3	2552	6	ABU33748	Protein e
403	29	76.3	670	4	AAU060913	Propionib	Aau060913	29	76.3	2572	6	ABU12083	Human NOV
404	29	76.3	670	6	ABW57432	Propionib	Abw57432	29	76.3	4472	2	AAW97245	Virulence
405	29	76.3	713	5	ABP65832	Bifidobac	Abp65832	29	76.3	4495	6	ABU69135	Human NOV
406	29	76.3	723	7	ADC01430	Enteroha	Adc01430	29	76.3	4529	5	AAU81016	Mouse alp
407	29	76.3	741	4	ABU95002	Human pro	Abu95002	29	76.3	4544	2	AAW47861	Alpha 2-M
408	29	76.3	741	6	ABU12082	Human NOV	Abu12082	29	76.3	4544	2	AAW60517	Human alp
409	29	76.3	743	4	AAU60990	Propionib	Aau60990	29	76.3	4544	4	AAW79091	Human pro
410	29	76.3	743	6	ABM57509	Propionib	Abm57509	29	76.3	4544	5	AAU81019	Human alp
411	29	76.3	744	4	AAW67243	Amino aci	Aaw67243	29	76.3	4544	6	ABP56839	Human LRP
412	29	76.3	780	4	AAW67244	Amino aci	Aaw67244	29	76.3	4544	6	ABP56839	Human LRP
413	29	76.3	800	4	AAU18148	Novel hum	Aau18148	29	76.3	4544	6	ABP56839	Human LRP
414	29	76.3	800	4	AAU17031	Human nov	Aau17031	29	76.3	4545	5	AAU74797	Mouse alp
415	29	76.3	800	4	ABU10492	Human cdn	Abu10492	29	76.3	4563	4	ABU11353	Human LDL
416	29	76.3	800	5	ABJ05775	Novel hum	Abj05775	29	76.3	5635	5	ABP60991	Novel hum
417	29	76.3	800	5	ABP67079	Human pol	Abp67079	29	76.3	10421	6	ABU33707	Protein e
418	29	76.3	820	6	ABM68831	Photorhab	Abm68831	29	76.3	10421	6	ABU33707	Protein e
419	29	76.3	822	4	ABG22519	Novel hum	Abg22519	29	76.3	6	AAW28912	Opioind pe	
420	29	76.3	832	5	ABG22519	Herbicida	Abg22519	29	76.3	6	AAW28912	Opioind pe	
421	29	76.3	833	5	ABG22785	Herbicida	Abg22785	29	76.3	6	AAW28912	Opioind pe	
422	29	76.3	845	2	AAW17890	Photorhab	Aaw17890	29	76.3	6	AAW28912	Opioind pe	
423	29	76.3	845	2	AAW56579	Fragment	Aaw56579	29	76.3	6	AAW28912	Opioind pe	
424	29	76.3	877	4	AAU16959	Human nov	Aau16959	29	76.3	6	AAW28912	Opioind pe	
425	29	76.3	883	4	ABG29974	Novel hum	Abg29974	29	76.3	6	AAW28912	Opioind pe	
426	29	76.3	962	5	ABP60992	Novel hum	Abp60992	29	76.3	6	AAW28912	Opioind pe	
427	29	76.3	974	5	ABG61872	Prostate	Abg61872	29	76.3	6	AAW28912	Opioind pe	
428	29	76.3	1001	7	ADP07847	Novel pro	Adp07847	29	76.3	6	AAW28912	Opioind pe	
429	29	76.3	1014	3	ABJ15909	E. coli p	Abj15909	29	76.3	6	AAW28912	Opioind pe	
430	29	76.3	1036	5	ABG70782	A. thalia	Abg70782	29	76.3	6	AAW28912	Opioind pe	
431	29	76.3	1036	6	AAE33693	Arabidops	Aae33693	29	76.3	6	AAW28912	Opioind pe	
432	29	76.3	1065	4	ABG25035	Novel hum	Abg25035	29	76.3	6	AAW28912	Opioind pe	
433	29	76.3	1068	4	AAW67244	Amino aci	Aaw67244	29	76.3	6	AAW28912	Opioind pe	
434	29	76.3	1092	5	ABG91150	Herbicida	Abg91150	29	76.3	6	AAW28912	Opioind pe	
435	29	76.3	1092	5	AAO22567	Wooden le	Aao22567	29	76.3	6	AAW28912	Opioind pe	
436	29	76.3	1148	2	AAW87895	Human JAG	Aaw87895	29	76.3	6	AAW28912	Opioind pe	
437	29	76.3	1160	6	ABU33712	Protein e	Abu33712	29	76.3	6	AAW28912	Opioind pe	
438	29	76.3	1213	4	ABG29914	Novel hum	Abg29914	29	76.3	6	AAW28912	Opioind pe	
439	29	76.3	1240	4	ABG03933	Novel hum	Abg03933	29	76.3	6	AAW28912	Opioind pe	
440	29	76.3	1251	5	ABG30538	Alpha-iso	Abg30538	29	76.3	6	AAW28912	Opioind pe	
441	29	76.3	1285	1	AAW33089	Pasteurel	Aaw33089	29	76.3	6	AAW28912	Opioind pe	
442	29	76.3	1286	5	ABG30564	Alpha-iso	Abg30564	29	76.3	6	AAW28912	Opioind pe	
443	29	76.3	1286	6	ADA26481	Alpha-iso	Ada26481	29	76.3	6	AAW28912	Opioind pe	
444	29	76.3	1293	4	ABBE5862	Drosophil	Abbe5862	29	76.3	6	AAW28912	Opioind pe	
445	29	76.3	1316	4	AAW39108	Human pol	Aaw39108	29	76.3	6	AAW28912	Opioind pe	
446	29	76.3	1336	5	AAW47771	Human thr	Aaw47771	29	76.3	6	AAW28912	Opioind pe	
447	29	76.3	1359	6	ABP57744	Human thr	Abp57744	29	76.3	6	AAW28912	Opioind pe	
448	29	76.3	1387	7	ADC39234	Novel hum	Adc39234	29	76.3	6	AAW28912	Opioind pe	
449	29	76.3	1401	4	AAW40894	Human pol	Aaw40894	29	76.3	6	AAW28912	Opioind pe	
450	29	76.3	1402	7	ADC39236	Novel hum	Adc39236	29	76.3	6	AAW28912	Opioind pe	
451	29	76.3	1422	2	AAW2067	Hepatitis	Aaw2067	29	76.3	6	AAW28912	Opioind pe	
452	29	76.3	1422	3	AAW09037	Hepatitis	Aaw09037	29	76.3	6	AAW28912	Opioind pe	
453	29	76.3	1576	6	ABP56835	Human MEG	Abp56835	29	76.3	6	AAW28912	Opioind pe	
454	29	76.3	1596	4	AAU28354	Novel hum	Aau28354	29	76.3	6	AAW28912	Opioind pe	
455	29	76.3	1630	6	AAE30206	Human LP2	Aae30206	29	76.3	6	AAW28912	Opioind pe	
456	29	76.3	1637	4	AAU28166	Novel hum	Aau28166	29	76.3	6	AAW28912	Opioind pe	
457	29	76.3	1641	4	ABU18356	Novel hum	Abu18356	29	76.3	6	AAW28912	Opioind pe	
458	29	76.3	1708	7	ADB90626	Mouse gua	Adb90626	29	76.3	6	AAW28912	Opioind pe	
459	29	76.3	1849	2	AAW17900	Photorhab	Aaw17900	29	76.3	6	AAW28912	Opioind pe	
460	29	76.3	1849	2	AAW56573	Toxin Tcd	Aaw56573	29	76.3	6	AAW28912	Opioind pe	
461	29	76.3	1902	6	ABU12084	Human NOV	Abu12084	29	76.3	6	AAW28912	Opioind pe	
462	29	76.3	1905	6	AAE29923	Human LP2	Aae29923	29	76.3	6	AAW28912	Opioind pe	
463	29	76.3	1906	7	ADD93399	Human lip	Add93399	29	76.3	6	AAW28912	Opioind pe	

ALIGNMENTS

RESULT 1
AAE21233
ID AAE21233 standard; protein; 53 AA.
XX
AC AAE21233;
DT 01-JUL-2002 (first entry)
XX
DE Human gene 18 encoded secreted protein HNNM45, SEQ ID NO:98.
XX
KW Human; secreted protein; immune disorder; antiallergic; antirheumatic;
KW rheumatoid arthritis; breast neoplasia; breast cancer; antiarthritic;
KW neurological disease; Alzheimer's disease; Parkinson's disease; trauma;
KW Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania;
KW antiinflammatory; ophthalmological; dermatological; immunostimulatory;
KW immunomodulatory; immunosuppressive; antibacterial; antiparasitic;
KW gene therapy; autoimmune disease; Huntington's disease; meningitis;
KW demyelinating disease; peripheral neuropathy; congenital malformation;
KW spinal cord injury; peripheral neuropathy; ischaemia; perception;
KW multiple sclerosis; infarction; haemorrhage; schizophrenia; dementia;
KW depression; panic disorder; learning disability; ALS; feeding disorder;
KW hyperproliferative disorder; sleep pattern; cardiovascular disorder;
KW reproductive disorder; digestive system disorder; behavioural disorder.
XX Homo sapiens.
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX
XX

ALIGNMENTS

RESULT 1

AAE21233
ID AAE21233 standard; protein; 53 AA.

XX AAE21233;

XX AC AAE21233;

XX DT 01-JUL-2002 (first entry)

XX DE Human gene 18 encoded secreted protein HNNM45, SEQ ID NO:98.

XX KW Human; secreted protein; immune disorder; antiallergic; antirheumatic; rheumatoid arthritis; breast neoplasia; breast cancer; antiarthritic; neurological disease; Alzheimer's disease; Parkinson's disease; trauma; Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania; antiinflammatory; ophthalmological; dermatological; immunostimulatory; immunomodulatory; immunosuppressive; antibacterial; antiporotic; gene therapy; autoimmune disease; Huntington's disease; meningitis; demyelinating disease; peripheral neuropathy; congenital malformation; spinal cord injury; peripheral neuropathy; ischaemia; perception; multiple sclerosis; infarction; haemorrhage; schizophrenia; dementia; depression; panic disorder; learning disability; AIDS; feeding disorder; hyperproliferative disorder; sleep pattern; cardiovascular disorder; reproductive disorder; digestive system disorder; behavioural disorder.

OS Homo sapiens.

XX Key Location/Qualifiers
XX Peptide 1..31
XX /label= Signal_peptide
XX Protein 32..53
XX /label= Mature_secreted_protein

```

FT Misc-difference 52
FT /label= Unknown
FT /note= "Encoded by NCA"
XX WO200216390-A1.
XX 28-FEB-2002.
XX 17-JAN-2001; 2001WO-US001435.
XX 18-AUG-2000; 2000US-0226282P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;
XX Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
XX Fiscella M, Ni J;
XX WPI; 2002-3041113/34.
XX N-PSDB; AAD33734.
XX An isolated nucleic acid molecule (I) comprising a polynucleotide which
XX encodes a polypeptide useful in the diagnosis and treatment of disorders
XX e.g. immune disorders.
XX Claim 11; Page 504; 504pp; English.
XX AAD33692-AAD33736 represent cDNAs corresponding to 21 human secreted
XX protein genes, and AAE21191-AAE21235 represent the proteins they encode.
XX AAE21236-AAE21280 represent human secreted protein fragments. The genes
XX and their corresponding secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 21 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of immune
XX or autoimmune diseases e.g. AIDS (acquired immune deficiency syndrome),
XX asthma, anaemia and rheumatoid arthritis, breast neoplasia and breast
XX cancer, neurological diseases e.g. Alzheimer's disease, Parkinson's
XX disease, Huntington's disease, Tourette syndrome, meningitis,
XX demyelinating disease, peripheral neuropathies, neoplasia, trauma,
XX congenital malformations, spinal cord injuries, toxic neuropathies
XX induced by neurotoxins, peripheral neuropathies, multiple sclerosis,
XX ischaemia and infarction, haemorrhages, schizophrenia, mania, dementia,
XX behaviours e.g. disorders in feeding, sleep patterns, balance and
XX perception, encephalitis, disorders in cardiovascular, neural/ sensory,
XX reproductive and digestive systems, behavioural disorders and
XX hyperproliferative disorder. The present sequence represents a human
XX secreted protein of the invention
XX SQ Sequence 53 AA;
XX Query Match 94.7%; Score 36; DB 5; Length 53;
XX Best Local Similarity 66.7%; Pred. No. 38;
XX Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX QY 1 WXXWHF 6
XX | |||
XX 21 WASWHF 26
XX Db
XX RESULT 2
XX ABG64892
XX ID ABG64892 standard; protein; 53 AA.
XX AC ABG64892;
XX XX
XX DT 27-AUG-2002 (first entry)
XX DE Human albumin fusion protein #1567.
XX XX

KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; anti-fertility; anti-inflammatory; anti-ulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX XX
OS Homo sapiens.
OS Synthetic.
XX WO20017137-A1.
XX 18-OCT-2001.
XX 12-APR-2001; 2001WO-US011988.
XX 12-APR-2000; 2000US-0229358P.
XX 25-APR-2000; 2000US-0199384P.
XX 21-DEC-2000; 2000US-0256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Haseltine WA;
XX WPI; 2002-010886/01.
XX New fusion protein for treating disease e.g. diabetes comprises an
XX albumin fused to a therapeutic protein.
XX Claim 1; Page 1582; 2102pp; English.
XX The present invention relates to albumin fusion proteins comprising a
XX therapeutic protein X and human albumin (HA, also known as human serum
XX albumin, HSA). The proteins are useful for treating a disease or disorder
XX that may be modulated by therapeutic protein X. The albumin extends the
XX shelf-life of protein X, and may increase its biological in vitro/in vivo
XX activity. The protein is useful for treating and diagnosing disorders
XX such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
XX disease, ulcerative colitis), immune disorders (e.g. acquired
XX immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
XX haematopoietic disorders, neural disorders (e.g. Alzheimer's,
XX Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
XX schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
XX ABG63326-ABG65518 represent albumin fusion proteins of the invention
XX SQ Sequence 53 AA;
XX Query Match 94.7%; Score 36; DB 5; Length 53;
XX Best Local Similarity 66.7%; Pred. No. 38;
XX Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX QY 1 WXXWHF 6
XX | |||
XX 21 WASWHF 26
XX Db
XX RESULT 3
XX AAB01505
XX ID AAB01505 standard; peptide; 6 AA.
XX AC AAB01505;
XX XX
XX DT 08-NOV-2000 (first entry)
XX DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX DE DNA binding; transcription factor; E2F; E2F-1; cell cycle; Dp-1;
XX KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
XX KW restenosis.
XX XX
OS Synthetic.

```



```

XX WO200044771-A1.
XX PS
XX 03-AUG-2000.
XX CC
XX 26-JAN-2000; 2000WO-GB000227.
XX PF
XX 26-JAN-1999; 99GB-00001710.
XX PR
XX (PROL-) PROLIFIX LTD.
XX PA
XX Mueller R, Kontermann RE, Montigiani S;
XX PI
XX WPI; 2000-532806/48.
XX DR
XX Peptides binding to the DNA binding domain of transcription factor E2F
XX PT and inhibiting cell cycle progression, useful for the treatment of
XX PT cancer.
XX PS
XX Example; Page 26; 42pp; English.
XX CC
XX Peptides which bind to the DNA binding domain of transcription factor E2F
XX CC and inhibit cell cycle progression may be useful as research agents to
XX CC investigate the interaction between E2F and DP-1, or the activation of
XX CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
XX CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
XX CC treatment of cancer or other proliferative disorders such as psoriasis
XX CC and restenosis.
XX PA
XX Sequence 6 AA;
XX SQ
XX Query Match 92.1%; Score 35; DB 3; Length 6;
XX Best Local Similarity 66.7%; Pred. NO. 1.4e+06;
XX Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX QY 1 WXXWHF 6
XX Db 1 WVAWHF 6
XX RESULT 5
XX ABB40893
XX ID ABB40893 standard; peptide; 36 AA.
XX XX
XX AC ABB40893;
XX XX
XX DT 04-FEB-2002 (first entry)
XX DE
XX DE Peptide #8399 encoded by human foetal liver single exon probe.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200157277-A2.
XX PD
XX PD 09-AUG-2001.
XX XX
XX PF 30-JAN-2001; 2001WO-US000669.
XX XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI
XX WPI; 2001-483447/52.
XX DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human fetal liver.
XX XX
XX Claim 27; SEQ ID NO 33528; 639pp + Sequence Listing; English.
XX XX
XX The invention relates to a single exon nucleic acid probe for measuring
XX CC human gene expression in a sample derived from human foetal liver. The
XX CC single exon nucleic acid probes may be used for predicting, measuring and
XX CC displaying gene expression in samples derived from human fetal liver. The
XX CC present sequence is a peptide encoded by a single exon nucleic acid probe
XX CC of the invention. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Query Match 92.1%; Score 35; DB 4; Length 36;
XX

```

```

XX WO200044771-A1.
XX PS
XX 03-AUG-2000.
XX CC
XX 26-JAN-2000; 2000WO-GB000227.
XX PF
XX 26-JAN-1999; 99GB-00001710.
XX PR
XX (PROL-) PROLIFIX LTD.
XX PA
XX Mueller R, Kontermann RE, Montigiani S;
XX PI
XX WPI; 2000-532806/48.
XX DR
XX Peptides binding to the DNA binding domain of transcription factor E2F
XX PT and inhibiting cell cycle progression, useful for the treatment of
XX PT cancer.
XX PS
XX Example; Page 26; 42pp; English.
XX CC
XX Peptides which bind to the DNA binding domain of transcription factor E2F
XX CC and inhibit cell cycle progression may be useful as research agents to
XX CC investigate the interaction between E2F and DP-1, or the activation of
XX CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
XX CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
XX CC treatment of cancer or other proliferative disorders such as psoriasis
XX CC and restenosis.
XX PA
XX Sequence 6 AA;
XX SQ
XX Query Match 92.1%; Score 35; DB 3; Length 6;
XX Best Local Similarity 66.7%; Pred. NO. 1.4e+06;
XX Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX QY 1 WXXWHF 6
XX Db 1 WVAWHF 6
XX RESULT 4
XX AAB01506
XX ID AAB01506 standard; peptide; 6 AA.
XX XX
XX AC AAB01506;
XX XX
XX DT 08-NOV-2000 (first entry)
XX DE
XX DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
XX KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
XX KW restenosis.
XX XX
XX OS Synthetic.
XX XX
XX WO200044771-A1.
XX PI
XX 03-AUG-2000.
XX XX
XX 26-JAN-2000; 2000WO-GB000227.
XX XX
XX 26-JAN-1999; 99GB-00001710.
XX XX
XX (PROL-) PROLIFIX LTD.
XX PA
XX Mueller R, Kontermann RE, Montigiani S;
XX PI
XX WPI; 2000-532806/48.
XX DR
XX Peptides binding to the DNA binding domain of transcription factor E2F
XX PT and inhibiting cell cycle progression, useful for the treatment of
XX PT cancer.
XX PT

```


PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
PT
XX Example 4; SEQ ID NO 33858; 650pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
XX Sequence 36 AA;
SQ
Query Match 92.1%; Score 35; DB 4; Length 36;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
DB 19 WRAWHF 24
RESULT 9
ABG56337
ID ABG56337 standard; peptide; 36 AA.
AC ABG56337;
XX
XX 25-FEB-2003 (first entry)
DT
XX Human liver peptide, SEQ ID NO 34985.
DE
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
XX Homo sapiens.
OS
XX WO200157273-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000664.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human adult liver.
XX
XX Claim 27; SEQ ID NO 34985; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridizes at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 36 AA;
SQ
Query Match 92.1%; Score 35; DB 4; Length 36;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
DB 19 WRAWHF 24
RESULT 10
ABP07322
ID ABP07322 standard; protein; 65 AA.
XX
XX ABP07322;
AC
XX 24-JUN-2002 (first entry)
DT
XX Human ORFX protein sequence SEQ ID NO:14626.
DE
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
XX Homo sapiens.
OS
XX WO200192523-A2.
PN
XX 06-DEC-2001.
PD
XX 29-MAY-2001; 2001WO-US010836.
PF
XX 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
PR
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach MD;
PI
XX WPI; 2002-106308/14.
DR N-PSDB; ABN23074.
DR
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
XX Disclosure; SEQ ID NO 14626; 1037pp; English.
PS
XX

CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). AHN15762 to AHN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. the sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 CC SQ Sequence 65 AA;

Query Match 92.1%; Score 35; DB 5; Length 65;
 Best Local Similarity 66.7%; Pred. No. 65;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 27 WKSWHF 32

RESULT 11
 ADA55310
 ID ADA55310 standard; protein; 113 AA.

AC ADA55310;

XX 20-NOV-2003 (first entry)

XX Human protein, SEQ ID 2878.

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease.

XX Homo sapiens.

XX EPI293569-A2.

XX 19-MAR-2003.

XX 21-MAR-2002; 2002EP-00006586.

XX 14-SEP-2001; 2001JP-00328381.

XX 24-JAN-2002; 2002US-0350435P.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-395539/38.

DR N-PSDB; ADA53671.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in

PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 14; SEQ ID NO 2878; 205pp; English.

XX The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

XX SQ Sequence 113 AA;

Query Match 92.1%; Score 35; DB 6; Length 113;

Best Local Similarity 66.7%; Pred. No. 11e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 71 WNTWHF 76

RESULT 12

ABB03211

ID ABB03211 standard; protein; 175 AA.

XX ABB03211;

XX 08-JAN-2002 (first entry)

XX Human musculoskeletal system related polypeptide SEQ ID NO 1158.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein;
 KW musculoskeletal system.

XX Homo sapiens.

XX WO200155367-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001338.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218230P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225286P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229503P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0229437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232633P.
PR 14-SEP-2000; 2000US-0232634P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234397P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251983P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-451937/48.
N-PSDB; AAL34793.
Isolated polypeptide for treating, preventing and/ or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis.
Claim 11; SEQ ID NO 1158; 781pp + Sequence Listing; English.
The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Sequence 175 AA;
Query Match 92.1%; Score 35; DB 4; Length 175;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-128199/12.
DR N-PSDB; ABX57781.
XX
PT Isolated nucleic acid molecules encoding musculoskeletal system
PR associated polypeptides, useful for detecting disorders, e.g. cancer.
XX
XX Claim 11; SEQ ID NO 1158; 321pp; English.
XX
CC The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals or
CC humans. The nucleic acid stimulates re-vascularisation of ischaemic
CC tissues associated with conditions such as thrombosis, arteriosclerosis,
CC and other cardiovascular conditions; treats wounds due to injuries,
CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
CC and limb regeneration; stimulates neuronal growth; can treat and prevent
CC neuronal damage occurring in certain disorders or neurodegenerative
CC conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-
CC related complex; stimulates chondrocyte growth, thus they can be used to
CC enhance bone and periodontal regeneration and aid in tissue transports or
CC bone grafts; prevents skin aging due to sunburn by stimulating
CC keratinocyte growth; prevents hair loss, since FGF family members

CC activate hair-forming cells and promotes melanocyte growth; stimulates
CC growth and differentiation of hematopoietic cells and bone marrow cells
CC when used in combination with other cytokines; maintains organs before
CC transplantation or for supporting cell culture of primary tissues;
CC induces tissue of mesodermal origin to differentiate in early embryos;
CC increases or decreases the differentiation or proliferation of embryonic
CC stem cells, besides, haematopoietic lineage; modulates mammalian
CC characteristics, such as, body height, weight, hair colour, eye colour,
CC skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,
CC cosmetic surgery); modulates mammalian metabolism; changes mammal's metal
CC state or physical state by influencing biorhythms, circadian rhythms,
CC depression, tendency for violence, tolerance for pain, reproductive
CC capabilities, hormonal or endocrine levels, appetite, libido, memory, or
CC stress; increases or decreases storage capabilities, fat content, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
CC components. This is the amino acid sequence of a novel human
CC musculoskeletal system antigen. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140
XX
SQ Sequence 175 AA;
Query Match 92.1%; Score 35; DB 6; Length 175;
Best Local Similarity 66.7%; Pred. NO. 1.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
Db 66 WGAMHF 71
RESULT 15
AAG09463
ID AAG09463 standard; protein; 256 AA.
XX
AC AAG09463;
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7409.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.

PR	21-OCT-1999;	99US-0160814P.	PR	14-MAY-1999;	99US-0134370P.
PR	21-OCT-1999;	99US-0160815P.	PR	18-MAY-1999;	99US-0134768P.
PR	22-OCT-1999;	99US-0160980P.	PR	19-MAY-1999;	99US-0134941P.
PR	22-OCT-1999;	99US-0160981P.	PR	20-MAY-1999;	99US-0135124P.
PR	22-OCT-1999;	99US-0160989P.	PR	21-MAY-1999;	99US-0135353P.
PR	23-OCT-1999;	99US-0161404P.	PR	24-MAY-1999;	99US-0135629P.
PR	25-OCT-1999;	99US-0161405P.	PR	25-MAY-1999;	99US-0136021P.
PR	25-OCT-1999;	99US-0161406P.	PR	27-MAY-1999;	99US-0136392P.
PR	26-OCT-1999;	99US-0161359P.	PR	28-MAY-1999;	99US-0136782P.
PR	26-OCT-1999;	99US-0161360P.	PR	01-JUN-1999;	99US-0137222P.
PR	26-OCT-1999;	99US-0161361P.	PR	03-JUN-1999;	99US-0137528P.
PR	28-OCT-1999;	99US-0161920P.	PR	04-JUN-1999;	99US-0137502P.
PR	28-OCT-1999;	99US-0161992P.	PR	07-JUN-1999;	99US-0137724P.
PR	28-OCT-1999;	99US-0161993P.	PR	08-JUN-1999;	99US-0138094P.
PR	28-OCT-1999;	99US-0162142P.	PR	10-JUN-1999;	99US-0138540P.
PR	28-OCT-1999;	99US-0162142P.	PR	10-JUN-1999;	99US-0138847P.
PR	28-OCT-1999;	99US-0162142P.	PR	14-JUN-1999;	99US-0139119P.
PR	28-OCT-1999;	99US-0162142P.	PR	16-JUN-1999;	99US-0139452P.
PR	28-OCT-1999;	99US-0162142P.	PR	16-JUN-1999;	99US-0139453P.
PR	28-OCT-1999;	99US-0162142P.	PR	17-JUN-1999;	99US-0139492P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139454P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139455P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139456P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139457P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139458P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139459P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139460P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139461P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139462P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139463P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139750P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139751P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139763P.
PR	28-OCT-1999;	99US-0162142P.	PR	21-JUN-1999;	99US-0139817P.
PR	28-OCT-1999;	99US-0162142P.	PR	22-JUN-1999;	99US-0139899P.
PR	28-OCT-1999;	99US-0162142P.	PR	23-JUN-1999;	99US-0140353P.
PR	28-OCT-1999;	99US-0162142P.	PR	23-JUN-1999;	99US-0140354P.
PR	28-OCT-1999;	99US-0162142P.	PR	24-JUN-1999;	99US-0140695P.
PR	28-OCT-1999;	99US-0162142P.	PR	28-JUN-1999;	99US-0140829P.
PR	28-OCT-1999;	99US-0162142P.	PR	29-JUN-1999;	99US-0140991P.
PR	28-OCT-1999;	99US-0162142P.	PR	30-JUN-1999;	99US-0141287P.
PR	28-OCT-1999;	99US-0162142P.	PR	01-JUL-1999;	99US-0141842P.
PR	28-OCT-1999;	99US-0162142P.	PR	01-JUL-1999;	99US-0142154P.
PR	28-OCT-1999;	99US-0162142P.	PR	02-JUL-1999;	99US-0142055P.
PR	28-OCT-1999;	99US-0162142P.	PR	06-JUL-1999;	99US-0142390P.
PR	28-OCT-1999;	99US-0162142P.	PR	08-JUL-1999;	99US-0142803P.
PR	28-OCT-1999;	99US-0162142P.	PR	09-JUL-1999;	99US-0142920P.
PR	28-OCT-1999;	99US-0162142P.	PR	12-JUL-1999;	99US-0142977P.
PR	28-OCT-1999;	99US-0162142P.	PR	13-JUL-1999;	99US-0143542P.
PR	28-OCT-1999;	99US-0162142P.	PR	14-JUL-1999;	99US-0143624P.
PR	28-OCT-1999;	99US-0162142P.	PR	15-JUL-1999;	99US-0144005P.
PR	28-OCT-1999;	99US-0162142P.	PR	16-JUL-1999;	99US-0144085P.
PR	28-OCT-1999;	99US-0162142P.			

```
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.

PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.1%; Score 35; DB 3; Length 256;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 72 WARWHF 77

RESULT 17
AAG48293
ID AAG48293 standard; protein; 256 AA.
XX
AC AAG48293;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60969.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
```

PR 27-MAY-1999; 99US-01336382P.
PR 28-MAY-1999; 99US-01336782P.
PR 01-JUN-1999; 99US-01337222P.
PR 03-JUN-1999; 99US-01337528P.
PR 04-JUN-1999; 99US-01337502P.
PR 07-JUN-1999; 99US-01337724P.
PR 08-JUN-1999; 99US-01338094P.
PR 10-JUN-1999; 99US-01338540P.
PR 10-JUN-1999; 99US-01338847P.
PR 14-JUN-1999; 99US-01339119P.
PR 16-JUN-1999; 99US-01339452P.
PR 16-JUN-1999; 99US-01339453P.
PR 17-JUN-1999; 99US-01339452P.
PR 18-JUN-1999; 99US-01339454P.
PR 18-JUN-1999; 99US-01339455P.
PR 18-JUN-1999; 99US-01339456P.
PR 18-JUN-1999; 99US-01339457P.
PR 18-JUN-1999; 99US-01339458P.
PR 18-JUN-1999; 99US-01339459P.
PR 18-JUN-1999; 99US-01339460P.
PR 18-JUN-1999; 99US-01339461P.
PR 18-JUN-1999; 99US-01339462P.
PR 18-JUN-1999; 99US-01339463P.
PR 18-JUN-1999; 99US-01339464P.
PR 18-JUN-1999; 99US-01339750P.
PR 18-JUN-1999; 99US-01339751P.
PR 21-JUN-1999; 99US-01339817P.
PR 22-JUN-1999; 99US-01339859P.
PR 23-JUN-1999; 99US-01403353P.
PR 23-JUN-1999; 99US-01403354P.
PR 24-JUN-1999; 99US-01403695P.
PR 28-JUN-1999; 99US-01408833P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142339P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144335P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144331P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 15-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.

```
PR 29-OCT-1999; 99US-0162142P.
Query Match 92.1%; Score 35; DB 3; Length 256;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
DB 72 WARWHF 77

RESULT 18
AAG93113
ID AAG93113 standard; protein; 258 AA.
XX AC AAG93113;
XX DT 26-SEP-2001 (first entry)
XX DE C glutamicum protein fragment SEQ ID NO: 6867.
XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX KW organic acid synthesis.
XX OS Corynebacterium glutamicum.
XX PN EP1108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-00127688.
XX FF 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-00159162.
XX PR 03-AUG-2000; 2000JP-00280988.
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX NAKAGAWA S, MIZOQUCHI H, ANDO S, HAYASHI M, OCHIAI K, YOKOI H;
XX TATEISHI N, SENOH A, IKEDA M, OZAKI A;
XX WPI; 2001-376931/40.
XX DR N-PSDB; AAH68332.
XX DR Novel polynucleotides derived from Corynebacterium bacteria, for identifying
XX PT mutation point of a gene, measuring expression of a gene, analyzing
XX PT expression profile or pattern of a gene and identifying homologous gene.
XX PS Claim 17; SEQ ID NO 6867; 246pp + Sequence Listing; English.
XX CC The present invention provides a number of nucleotide and protein
XX CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
XX CC are useful for identifying the mutation point of a gene derived from a
XX CC mutant of corynebacterium bacterium, measuring expression amount and analyzing
XX CC the expression profile or expression pattern of a gene derived from
XX CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
XX CC corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
XX CC acids, nucleic acids, vitamins, saccharides and organic acids,
XX CC particularly L-lysine. The present sequence is a protein described in the
XX CC exemplification of the invention. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from the European Patent Office
XX Sequence 258 AA;
Query Match 92.1%; Score 35; DB 4; Length 258;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
DB 16 WHAWHF 21

RESULT 19
AAG50316
ID AAG50316 standard; protein; 324 AA.
XX AC AAG50316;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 63749.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 01-APR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 28-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132484P.
XX PR 05-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 06-MAY-1999; 99US-0132487P.
XX PR 07-MAY-1999; 99US-0132863P.
XX PR 11-MAY-1999; 99US-0134258P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 14-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135533P.
XX PR 24-MAY-1999; 99US-0135629P.
XX PR 25-MAY-1999; 99US-0136021P.
XX PR 27-MAY-1999; 99US-0136392P.
XX PR 28-MAY-1999; 99US-0136782P.
XX PR 01-JUN-1999; 99US-0137222P.
XX PR 03-JUN-1999; 99US-0137528P.
XX PR 04-JUN-1999; 99US-0137502P.
XX PR 07-JUN-1999; 99US-0137724P.
XX PR 08-JUN-1999; 99US-0138094P.
XX PR 10-JUN-1999; 99US-0138540P.
XX PR 10-JUN-1999; 99US-0138847P.
XX PR 14-JUN-1999; 99US-0139119P.
XX PR 16-JUN-1999; 99US-0139452P.
XX PR 16-JUN-1999; 99US-0139453P.
XX PR 17-JUN-1999; 99US-0139492P.
XX PR 18-JUN-1999; 99US-0139454P.
XX PR 18-JUN-1999; 99US-0139455P.
XX PR 18-JUN-1999; 99US-0139456P.
XX PR 18-JUN-1999; 99US-0139457P.
```

PR	18-JUN-1999;	90US-0139458P
PR	18-JUN-1999;	90US-0139459P
PR	18-JUN-1999;	90US-0139460P
PR	18-JUN-1999;	90US-0139461P
PR	18-JUN-1999;	90US-0139462P
PR	18-JUN-1999;	90US-0139463P
PR	18-JUN-1999;	90US-0139750P
PR	18-JUN-1999;	90US-0139763P
PR	21-JUN-1999;	90US-0139817P
PR	22-JUN-1999;	90US-0139899P
PR	23-JUN-1999;	90US-0140353P
PR	23-JUN-1999;	90US-0140354P
PR	24-JUN-1999;	90US-0140695P
PR	24-JUN-1999;	90US-0140823P
PR	28-JUN-1999;	90US-0140991P
PR	30-JUN-1999;	90US-0141847P
PR	01-JUL-1999;	90US-0141848P
PR	01-JUL-1999;	90US-0142154P
PR	02-JUL-1999;	90US-0142055P
PR	06-JUL-1999;	90US-0142390P
PR	08-JUL-1999;	90US-0142803P
PR	09-JUL-1999;	90US-0142920P
PR	12-JUL-1999;	90US-0142977P
PR	13-JUL-1999;	90US-0143542P
PR	14-JUL-1999;	90US-0143624P
PR	15-JUL-1999;	90US-0144005P
PR	16-JUL-1999;	90US-0144085P
PR	16-JUL-1999;	90US-0144086P
PR	18-JUL-1999;	90US-0144325P
PR	19-JUL-1999;	90US-0144331P
PR	19-JUL-1999;	90US-0144332P
PR	19-JUL-1999;	90US-0144333P
PR	19-JUL-1999;	90US-0144334P
PR	19-JUL-1999;	90US-0144335P
PR	20-JUL-1999;	90US-0144352P
PR	20-JUL-1999;	90US-0144632P
PR	20-JUL-1999;	90US-0144884P
PR	21-JUL-1999;	90US-0144814P
PR	21-JUL-1999;	90US-0145086P
PR	21-JUL-1999;	90US-0145087P
PR	22-JUL-1999;	90US-0145085P
PR	22-JUL-1999;	90US-0145087P
PR	22-JUL-1999;	90US-0145089P
PR	22-JUL-1999;	90US-0145192P
PR	23-JUL-1999;	90US-0145145P
PR	23-JUL-1999;	90US-0145218P
PR	23-JUL-1999;	90US-0145224P
PR	26-JUL-1999;	90US-0145276P
PR	27-JUL-1999;	90US-0145913P
PR	27-JUL-1999;	90US-0145918P
PR	28-JUL-1999;	90US-0145919P
PR	28-JUL-1999;	90US-0145951P
PR	02-AUG-1999;	90US-0146386P
PR	02-AUG-1999;	90US-0146388P
PR	02-AUG-1999;	90US-0146389P
PR	03-AUG-1999;	90US-0147038P
PR	04-AUG-1999;	90US-0147204P
PR	05-AUG-1999;	90US-0147302P
PR	05-AUG-1999;	90US-0147192P
PR	06-AUG-1999;	90US-0147260P
PR	06-AUG-1999;	90US-0147303P
PR	06-AUG-1999;	90US-0147416P
PR	09-AUG-1999;	90US-0147493P
PR	09-AUG-1999;	90US-0147935P
PR	10-AUG-1999;	90US-0148171P
PR	11-AUG-1999;	90US-0148319P
PR	12-AUG-1999;	90US-0148341P
PR	13-AUG-1999;	90US-0148565P
PR	13-AUG-1999;	90US-0148684P
PR	16-AUG-1999;	90US-0149368P
PR	17-AUG-1999;	90US-0149175P
PR	18-AUG-1999;	90US-0149426P
PR	20-AUG-1999;	90US-0149722P


```
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154033P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155133P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.1%; Score 35; DB 3; Length 325;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 WXXWHF 6
Db 141 WARWHF 146

RESULT 21
AAG44074
ID AAG44074 standard; protein; 325 AA.
XX AC
XX AAG44074;
XX AC
DT 18-OCT-2000 (first entry)
XX AC
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55162.
XX AC
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
```

```
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD C6-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR C5-MAR-1999; 99US-0123180P.
PR C9-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR C6-APR-1999; 99US-0128234P.
PR C8-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
```

PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.

PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.1%; Score 35; DB 3; Length 325;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 141 WARWHF 146

RESULT 22

AAG09462
ID AAG09462 standard; protein; 325 AA.

XX AC AAG09462;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7408.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX OS Arabidopsis thaliana.

XX PN E21033405-A2.

XX PD 06-SEP-2000.

XX XX

PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 28-OCT-1999; 99US-0162142P.
 Query Match 92.1%; Score 35; DB 3; Length 325;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 DB 141 WARWHF 146
 RESULT 24
 ID ADB95074
 AC ADB95074;
 DT 04-DEC-2003 (first entry)
 DE A. thaliana protein 52949A #SEQ ID 72.
 KW Plant; herbicide; weed; crop field; growth; development.
 OS Arabidopsis thaliana.
 XX WO2003008440-A2.
 PN 30-JAN-2003.
 PD 16-JUL-2002; 2002WO-EP007929.
 PF 16-JUL-2001; 2001US-0305806P.
 PR 20-FEB-2002; 2002US-0358416P.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA Levin JZ, Patton DA, Mcelver JA, Budziszewski GJ, Zhou Q, Aux GW;
 PI Tossberg J, Wegrich Glover L, Ashby CS, Thomas CR, Madhavan E;
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 28-OCT-1999; 99US-0162142P.
 Query Match 92.1%; Score 35; DB 7; Length 325;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 DB 141 WARWHF 146
 RESULT 24
 ID AAW46287
 AC AAW46287;
 DT 17-OCT-2003 (revised)
 DT 12-AUG-1998 (first entry)
 DE Rhizobium meliloti DNA methyltransferase amino acid sequence.
 KW DNA methyltransferase; cell cycle regulated methyltransferase; CorM;
 KW cell viability; assay; activity; inhibitor; detection; antibiotic.
 OS Sinorhizobium meliloti.
 XX WO9812206-A1.
 PN 26-MAR-1998.
 PD 17-SEP-1997; 97WO-US016593.
 PF 19-SEP-1996; 96US-0020089P.
 PR (STRD) UNIV LELAND STANFORD JUNIOR.
 PA (PENN-) PENN STATE RES FOUND.
 XX Shapero L, Benkovic SJ, Wright R, Stephens C, Kahng LS, Berdis A;
 PI Lee I;
 XX

PI Lewis S, Dunn J, Cates E, Law MD;
 XX WPI; 2003-229557/22.
 DR N-PSDB; ADB95073.
 XX
 PT Identifying an herbicidal compound, useful for controlling undesirable
 PT vegetation, comprises combining a polypeptide with a compound to be
 PT tested for the ability to bind to the polypeptide or inhibit the activity
 PT of the polypeptide.
 XX
 XX Claim 4; SEQ ID NO 72; 273pp; English.
 PS
 CC The invention relates to a method for identifying a herbicidal compound.
 CC The method of the invention comprises combining a polypeptide having at
 CC least 90% identical to any one of 48 69-1008 residue amino acid sequences
 CC (designated as PI-P48), given in the specification, with a compound to be
 CC tested for the ability to bind to the polypeptide or inhibit the activity
 CC of the polypeptide, under conditions conducive to binding or inhibiting the
 CC respectively. Also disclosed is a method for killing or inhibiting the
 CC growth or viability of a plant by applying to the plant the herbicidal
 CC compound identified by the novel method, a chimeric construct comprising
 CC a promoter operatively linked to the nucleic acid molecule, a recombinant
 CC vector comprising the chimeric construct and a host cell comprising the
 CC nucleic acid molecule. The method and polypeptides are useful in
 CC screening assays to identify compounds that interact or inhibit the
 CC polypeptides, thus as potential herbicides to control undesirable
 CC vegetation such as weeds in crop fields. Nucleic acid molecules (odd
 CC numbers between ADB95003 and ADB95097) isolated from Arabidopsis thaliana
 CC comprising nucleotide sequences that encode proteins (even numbers
 CC between ADB95004-ADB95098) are essential for plant growth and
 CC development.
 XX
 XX Sequence 325 AA;

DR WPI: 1998-217199/19.
 DR N-PSDB; AAV26374.
 XX Bacterial methyl-transferase proteins - used to isolate antibiotics and
 PT inhibitors of bacterial growth.
 PT Claim 1; Fig 2; 66pp; English.
 XX
 CC The present sequence represents a DNA methyltransferase of Rhizobium
 CC meliloti. The methyltransferase is a cell cycle regulated
 CC methyltransferase (CcrM). It is essential for cell viability. The
 CC specification describes an assay for methyltransferase activity
 CC comprising contacting a processive methyltransferase with a substrate
 CC selected AAV26378-30, and further contacting the processive
 CC methyltransferase with a methyl donor (e.g. S-adenosyl methionine) prior
 CC to or at the same time as the addition of the DNA substrate, where the
 CC methyltransferase methylates the DNA substrate. The methyltransferase
 CC proteins can be used in an assay for screening for inhibitors of DNA
 CC methyltransferase activity. They can also be used in an assay for
 CC detecting antibiotics that target processive adenine methyltransferases.
 CC Inhibitors of the methyltransferase activity results in a migration or
 CC elimination of the subject bacteria to infect and or grow and proliferate
 CC in an animal or plant host. (Updated on 17-OCT-2003 to standardise OS
 CC field)
 XX Sequence 376 AA;
 SQ
 Query Match 92.1%; Score 35; DB 2; Length 376;
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 Db 347 WTFWHF 352
 RESULT 25
 AAW46288
 ID AAW46288 standard; protein; 377 AA.
 XX
 AC AAW46288;
 XX
 DT 17-OCT-2003 (revised)
 DT 12-AUG-1998 (first entry)
 XX
 DE Brucella abortus DNA methyltransferase.
 XX
 KW DNA methyltransferase; cell cycle regulated methyltransferase; CcrM;
 KW cell viability; assay; activity; inhibitor; detection; antibiotic.
 XX
 OS Brucella melitensis biovar Abortus.
 XX
 PN WO9812206-A1.
 XX
 PD 26-MAR-1998.
 XX
 PF 17-SEP-1997; 97WO-US016593.
 XX
 PR 19-SEP-1996; 96US-0020089P.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PA (PENN-) PENN STATE RES FOUND.
 XX
 XX Shapiro L, Benkovic SJ, Wright R, Stephens C, Kahng LS, Berdis A;
 PI Lee I;
 XX WPI; 1998-217199/19.
 DR N-PSDB; AAV26375.
 XX
 CC Bacterial methyl-transferase proteins - used to isolate antibiotics and
 PT inhibitors of bacterial growth.
 PT Claim 1; Fig 4; 66pp; English.

XX The present sequence represents a DNA methyltransferase of Brucella
 CC abortus. The methyltransferase is a cell cycle regulated
 CC methyltransferase (CcrM). It is essential for cell viability. The
 CC specification describes an assay for methyltransferase activity
 CC comprising contacting a processive methyltransferase with a substrate
 CC selected AAV26378-30, and further contacting the processive
 CC methyltransferase with a methyl donor (e.g. S-adenosyl methionine) prior
 CC to or at the same time as the addition of the DNA substrate, where the
 CC methyltransferase methylates the DNA substrate. The methyltransferase
 CC proteins can be used in an assay for screening for inhibitors of DNA
 CC methyltransferase activity. They can also be used in an assay for
 CC detecting antibiotics that target processive adenine methyltransferases.
 CC Inhibitors of the methyltransferase activity results in a migration or
 CC elimination of the subject bacteria to infect and or grow and proliferate
 CC in an animal or plant host. (Updated on 17-OCT-2003 to standardise OS
 CC field)
 XX Sequence 377 AA;
 SQ
 Query Match 92.1%; Score 35; DB 2; Length 377;
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 Db 347 WTFWHF 352
 RESULT 26
 AAG50315
 ID AAG50315 standard; protein; 423 AA.
 XX
 AC AAG50315;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 63748.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 04-MAY-1999; 99US-0132407P.
 PR 05-MAY-1999; 99US-0132484P.
 PR 06-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 07-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134212P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135623P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140901P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145122P.
PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151338P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0155659P.
PR 29-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.

```
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.1%; Score 35; DB 3; Length 423;
Best Local Similarity 66.7%; Pred. No. 3.Se+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 125 WYSWHF 130

RESULT 27
AAG44023
ID AAG44023 standard; protein; 452 AA.
XX
AC AAG44023;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55093.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-012348P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
```

PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147152P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.18; Score 35; DB 3; Length 452;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
| | | | |
Db 125 WYSWHF 130

RESULT 28
AAG50314
ID AAG50314 standard; protein; 493 AA.
XX AAG50314;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63747.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137722P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149317P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159253P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160844P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match

92.1%; Score 35; DB 3; Length 493;

Best Local Similarity 66.7%; Pred. No. 4e+02; Mismatches 0; Indels 2; Gaps 0;

Matches 4; Conservative 0; WXSXWHF 6

QY 1 WXSXWHF 6

Db 195 WXSXWHF 200

RESULT 29

AAG44022

ID AAG44022 standard; protein; 520 AA.

XX AC AAG44022;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 55092.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 04-MAY-1999; 99US-0132048P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 02-JUL-1999; 99US-0142154P.

PR 06-JUL-1999; 99US-0142055P.

PR 08-JUL-1999; 99US-0142390P.

PR 09-JUL-1999; 99US-0142803P.

PR 12-JUL-1999; 99US-0142920P.

PR 13-JUL-1999; 99US-0142977P.

PR 14-JUL-1999; 99US-0143542P.

PR 15-JUL-1999; 99US-0143624P.

PR 16-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 23-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 03-AUG-1999; 99US-0146389P.

PR 04-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154778P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161403P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.1%; Score 35; DB 3; Length 520;
Best Local Similarity 66.7%; Pred.No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 193 WYSWHF 198

RESULT 30
AAG44021
ID AAG44021 standard; protein; 522 AA.
XX
AC AAG44021;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55091.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140981P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147460P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.

PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.1%; Score 35; DB 3; Length 522;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 195 WYSWHF 200

RESULT 31
ABJ25385
ID ABJ25385 standard; protein; 522 AA.
XX
AC ABJ25385;
XX

DT 16-APR-2003 (first entry)
 XX Aspergillus fumigatus essential gene protein #43.
 DE Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
 KW cancer; contamination; biofilm; antibody; immune response.
 XX Aspergillus fumigatus.
 OS WO200286090-A2.
 XX 31-OCT-2002.
 XX 23-APR-2002; 2002WO-US013142.
 PF 23-APR-2001; 2001US-0285697P.
 PR 27-APR-2001; 2001US-0287066P.
 PR 05-JUN-2001; 2001US-0295890P.
 PR 09-JUL-2001; 2001US-0303899P.
 PR 31-AUG-2001; 2001US-0316362P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
 XX WPI; 2003-093124/08.
 DR New purified or isolated nucleic acids of essential genes of Aspergillus
 XX fumigatus, useful for treating or preventing infections by A. fumigatus,
 XX or for treating a non-infectious disease in a subject e.g. cancer.
 PS Disclosure; Page; 175pp; English.
 XX The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as A. fumigatus, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for
 CC expressing recombinant protein for characterisation, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of A.
 CC fumigatus to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This sequence represents a protein of one of the essential genes
 CC of Aspergillus fumigatus of the invention
 XX
 SQ Sequence 522 AA;

Query Match 92.1%; Score 35; DB 6; Length 522;
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 DB 198 WYSWHF 203

RESULT 32
 ABJ25985

ID ABJ25985 standard; protein; 522 AA.
 XX AC ABJ25985;
 XX 16-APR-2003 (first entry)
 DT Aspergillus fumigatus essential gene protein #643.
 DE Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
 KW cancer; contamination; biofilm; antibody; immune response.
 XX Aspergillus fumigatus.
 OS WO200286090-A2.
 XX 31-OCT-2002.
 XX 23-APR-2002; 2002WO-US013142.
 PF 23-APR-2001; 2001US-0285697P.
 PR 27-APR-2001; 2001US-0287066P.
 PR 05-JUN-2001; 2001US-0295890P.
 PR 09-JUL-2001; 2001US-0303899P.
 PR 31-AUG-2001; 2001US-0316362P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
 XX WPI; 2003-093124/08.
 DR New purified or isolated nucleic acids of essential genes of Aspergillus
 XX fumigatus, useful for treating or preventing infections by A. fumigatus,
 XX or for treating a non-infectious disease in a subject e.g. cancer.
 PS Disclosure; Page; 175pp; English.
 XX The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as A. fumigatus, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for
 CC expressing recombinant protein for characterisation, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of A.
 CC fumigatus to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This sequence represents a protein of one of the essential genes
 CC of Aspergillus fumigatus of the invention
 XX
 SQ Sequence 522 AA;

Query Match 92.1%; Score 35; DB 6; Length 522;
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 DB 198 WYSWHF 203

RESULT 33

AAB01499
ID AAB01499 standard; peptide; 6 AA.

XX AC AAB01499;
XX DT 08-NOV-2000 (first entry)
XX DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
XX KW restenosis.
XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 2 /note= "Any amino acid"

FT FT Misc-difference 3 /note= "Any amino acid"

FT FT

XX WO200044771-A1.

XX PD 03-AUG-2000.

XX PF 26-JAN-2000; 2000WO-GB000227.

XX PR 26-JAN-1999; 99GB-00001710.

XX PA (PROL-) PROLIFIX LTD.

XX PI Mueller R, Kontermann RE, Montigiani S;

XX DR WPI; 2000-532806/48.

XX PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX PS Claim 4; Page 9; 42pp; English.

XX CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis

XX SQ Sequence 6 AA;

Query Match 89.5%; Score 34; DB 3; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WXXWHF 6

DB 1 WXXWHF 6

RESULT 34

AAB01492
ID AAB01492 standard; peptide; 6 AA.

XX AC AAB01492;

XX DT 08-NOV-2000 (first entry)

XX DE Peptide which binds to transcription factor E2F-1 DNA binding domain.
XX

KW DNA binding; transcription factor; E2F; E2F-1; cell cycle; DP-1;
KW activation; transcription; apoptosis; proliferative disorder; psoriasis;
XX KW restenosis.

XX OS Synthetic.

XX PN WO200044771-A1.

XX PD 03-AUG-2000.

XX PF 26-JAN-2000; 2000WO-GB000227.

XX PR 26-JAN-1999; 99GB-00001710.

XX PA (PROL-) PROLIFIX LTD.

XX PI Mueller R, Kontermann RE, Montigiani S;

XX DR WPI; 2000-532806/48.

XX PT Peptides binding to the DNA binding domain of transcription factor E2F
PT and inhibiting cell cycle progression, useful for the treatment of
PT cancer.
XX PS Claim 6; Page 2; 42pp; English.

XX CC Peptides which bind to the DNA binding domain of transcription factor E2F
CC and inhibit cell cycle progression may be useful as research agents to
CC investigate the interaction between E2F and DP-1, or the activation of
CC transcription by E2F-1/DP-1 heterodimers. They may also be used for
CC inducing apoptosis and/or cell cycle arrest in a cell, particularly for
CC treatment of cancer or other proliferative disorders such as psoriasis
CC and restenosis

XX SQ Sequence 6 AA;

Query Match 89.5%; Score 34; DB 3; Length 6;

Best Local Similarity 66.7%; Pred. No. 1.4e+06;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WXXWHF 6

DB 1 WVRWHF 6

RESULT 35

AAU14683
ID AAU14683 standard; protein; 43 AA.

XX AC AAU14683;

XX DT 24-OCT-2001 (first entry)

XX DE Novel bone marrow polypeptide #82.

XX KW Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
KW haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
KW wound healing; nutritional supplement; immune disorder;
XX KW severe combined immunodeficiency; SCID.
XX OS Homo sapiens.
XX PN WO200157187-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US003782.

XX PR 03-FEB-2000; 2000US-00496914.

XX PR 20-JUN-2000; 2000US-00598075.

XX PR 19-JUL-2000; 2000US-00620325.

XX PR 30-NOV-2000; 2000US-0250683P.

(HYSE-) HYSEQ INC.
 Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
 Ren F, Drmanac RT;
 WPI; 2001-488875/53.
 N-PSDB; AAS22988.
 Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and gene therapy.
 Claim 10; Page 248; 392pp; English.
 AAU14602-AAU14794 represent novel bone marrow polypeptides of the invention. The proteins and corresponding coding sequences may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate bone marrow polypeptide expression. For example, to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of the polypeptides by expressing inactive proteins or to supplement the patient's own production of the polypeptides. Additionally, the nucleic acids may be used to produce the polypeptides, by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The nucleic acid and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and therefore which patients may be in need of restorative therapy. The proteins may also be used as antigens in the production of antibodies against bone marrow proteins and in assays to identify modulators of their expression and activity. The anti-bone marrow protein antibodies and antagonists may also be used to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of the protein in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The proteins may be used to regulate haematopoiesis activity, and consequently in the treatment of myeloid or lymph cell disorders; in tissue regeneration, such as wound healing; as a nutritional supplement; and in treatment of immune disorders such as severe combined immunodeficiency (SCID)

Sequence 43 AA;
 Query Match 89.5%; Score 34; DB 4; Length 43;
 Best Local Similarity 66.7%; Pred. No. 66;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 | | | |
 Db 14 WLIWHF 19

RESULT 36
 AAU14777
 ID AAU14777 standard; protein; 96 AA.
 AC AAU14777;
 DT 24-OCT-2001 (first entry)
 DE Novel bone marrow polypeptide #176.
 KW Bone marrow; diagnostic; therapeutic; gene therapy; antigenic; haematopoiesis; myeloid; lymph cell disorder; tissue regeneration; wound healing; nutritional supplement; immune disorder;
 KW severe combined immunodeficiency; SCID.
 OS Homo sapiens.
 XX WO200157187-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US003782.
 XX 03-FEB-2000; 2000US-00496914.

PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 30-NOV-2000; 2000US-0250683P.
 XX (HYSE-) HYSEQ INC.
 XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
 PI Ren F, Drmanac RT;
 XX WPI; 2001-488875/53.
 DR N-PSDB; AAS23082.
 XX Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and gene therapy.
 XX Claim 10; Page 137; 392pp; English.
 CC AAU14602-AAU14794 represent novel bone marrow polypeptides of the invention. The proteins and corresponding coding sequences may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate bone marrow polypeptide expression. For example, to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of the polypeptides by expressing inactive proteins or to supplement the patient's own production of the polypeptides. Additionally, the nucleic acids may be used to produce the polypeptides, by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The nucleic acid and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and therefore which patients may be in need of restorative therapy. The proteins may also be used as antigens in the production of antibodies against bone marrow proteins and in assays to identify modulators of their expression and activity. The anti-bone marrow protein antibodies and antagonists may also be used to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of the protein in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The proteins may be used to regulate haematopoiesis activity, and consequently in the treatment of myeloid or lymph cell disorders; in tissue regeneration, such as wound healing; as a nutritional supplement; and in treatment of immune disorders such as severe combined immunodeficiency (SCID)

Sequence 96 AA;
 Query Match 89.5%; Score 34; DB 4; Length 96;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 | | | |
 Db 67 WLIWHF 72

RESULT 37
 ABP06139
 ID ABP06139 standard; protein; 119 AA.
 XX ABP06139;
 AC ABP06139;
 XX 25-JUN-2002 (first entry)
 DT Human ORFX protein sequence SEQ ID NO:12260.
 DE Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis; hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; degenerative disorder; osteoarthritis; neurodegenerative disorder; cardiovascular disease; diabetes mellitus; systemic lupus erythematosus; hypertension; hypothyroidism; cholesterol ester storage disease; immune deficiency; immune disorder; infectious disease; autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis; myasthenia gravis.
 KW Homo sapiens.
 XX

XX WO200192523-A2.
PN
XX
PD 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US010836.
XX
XX 30-MAY-2000; 2000US-0206132P.
PR
XX 29-AUG-2000; 2000US-0228716P.
PR
XX (CURA-) CURAGEN CORP.
PA
XX Shimkets RA, Leach MD;
XX
XX WPI; 2002-106308/14.
DR
XX N-PSDB; ABN21891.
DR
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
PT
XX Disclosure; SEQ ID NO 12260; 1037pp; English.
PS
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 119 AA;
XX
Query Match 89.5%; Score 34; DB 5; Length 119;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
Db 65 WYGVWF 70
XX
RESULT 38
AAG83334
ID AAG83334 standard; protein; 173 AA.
XX
XX AAG83334;
AC
XX 06-SEP-2001 (first entry)
DT
XX P patens lipid metabolism related protein #52.
DE
XX Moss; LMRP; lipid metabolism related protein; polynunsaturated fatty acid;
KW fine chemical; transgenic plant.
KW
XX Physcomitrella patens.
OS

XX WO200138541-A1.
PN
XX 31-MAY-2001.
PD
XX 25-NOV-1999; 99WO-EP009108.
PF
XX 25-NOV-1999; 99WO-EP009108.
PR
XX (BADI) BASF PLANT SCI GMBH.
PA
XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;
PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;
XX
XX WPI; 2001-381293/40.
DR
XX New isolated nucleic acid molecule encoding Lipid Metabolism Related
PT Proteins useful in the production of fine chemicals.
PT
XX Claim 31; Page 105; 113pp; English.
PS
XX The present invention provides the protein and coding sequences of a
CC number of moss lipid metabolism related proteins (LMRPs). The moss
CC Physcomitrella patens is one of the few plants able to produce
CC polynunsaturated fatty acids, and the sequences can be used to create
CC transgenic plants also capable of producing them. They can also be used
CC to identify the presence of P. patens and in the production of fine
CC chemicals. The present sequence is one of the proteins of the invention
XX
SQ Sequence 173 AA;
XX
Query Match 89.5%; Score 34; DB 4; Length 173;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
Db 73 WLLWHF 78
XX
RESULT 39
AAG80884
ID AAG80884 standard; protein; 173 AA.
XX
XX AAG80884;
AC
XX 28-AUG-2001 (first entry)
DT
XX Lipid modification protein sequence #5.
DE
XX Moss; Physcomitrella patens; lipid metabolism related protein; LMRP;
KW lipid biosynthesis; lipid modification; lipid degradation; cofactor;
KW fatty acid transport; genetic engineering; fatty acid; enzyme; plant;
KW microorganism; polynunsaturated fatty acid; oilseed plant; maize; wheat;
KW biotic stress tolerance; abiotic stress tolerance; rye; oat; triticale;
KW rice; barley; soybean; peanut; cotton; rapeseed; canola; manihot; pepper;
KW sunflower; tagetes; potato; tobacco; eggplant; tomato; Vicia; pea;
KW alfalfa; coffee; cacao; tea; salix; oil palm; coconut; perennial grass;
KW forage crop.
XX
XX Physcomitrella patens.
OS
XX WO200138484-A2.
PN
XX 31-MAY-2001.
PD
XX 22-NOV-2000; 2000WO-EP011615.
PF
XX 25-NOV-1999; 99WO-EP009108.
PR
XX (BADI) BASF PLANT SCI GMBH.
PA
XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;
PI

```

PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;
XX WPI; 2001-367669/38.
XX
XX Nucleic acids encoding lipid metabolism related proteins from
PT Psychomitrella patens useful to produce fine chemicals in modified
PT organisms, particularly polyunsaturated fatty acids in oilseed plants.
XX
XX Claim 31; Page 114; 120pp; English.
XX
XX The present invention describes isolated nucleic acid sequences which
XX encode lipid metabolism related proteins (LMRP). The LMRP nucleic acids
XX can be used to modify lipids and fatty acids, cofactors and enzymes in
XX microorganisms and plants, particularly to produce polyunsaturated fatty
XX acids, and are especially useful in oilseed plants. The nucleic acids may
XX also confer biotic or abiotic stress tolerance, particularly to maize,
XX wheat, rye, oat, triticale, rice, barley, soybean, peanut, cotton,
XX rapeseed, canola, manihot, pepper, sunflower, tagetes, potato, tobacco,
XX eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao, tea, Salix
XX species, oil palm, coconut, perennial grasses and forage crops. AAH50878
XX to AAH50882 represent primers used in the exemplification of the present
XX invention. AAH50883 to AAH50988 represents LMRP nucleotide sequences, and
XX AAH50843 to AAH50928 represent LMRP protein sequences, given in the
XX present invention
XX
XX Sequence 173 AA;
XX
XX Query Match 89.5%; Score 34; DB 4; Length 173;
XX Best Local Similarity 66.7%; Pred. No. 2.3e+02;
XX Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 WXXWHF 6
XX | | | |
XX 73 WLLWHF 78
XX
XX RESULT 40
XX ADA54839
XX ID ADA54839 standard; protein; 190 AA.
XX
XX AC ADA54839;
XX
XX DT 20-NOV-2003 (first entry)
XX
XX DE Human protein, SEQ ID 2407.
XX
XX KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
XX Gene Therapy; human; secretory protein; membrane proteins; cancer;
XX inflammatory disease; osteoporosis; neurological disease.
XX
XX OS Homo sapiens.
XX
XX PN EP1293569-A2.
XX
XX XX EP1293569-A2.
XX
XX PD 19-MAR-2003.
XX
XX XX 21-MAR-2002; 2002EP-00006586.
XX
XX PF 14-SEP-2001; 2001JP-00328381.
XX
XX PR 24-JAN-2002; 2002US-0350433P.
XX
XX XX (HELI-) HELIX RES INST.
XX
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI; 2003-395539/38.
XX
XX DR N-PSDB; ADA53200.
XX
XX PT New polynucleotides encoding full-length polypeptides, e.g. secretory
XX and/or membrane proteins, useful for developing medicines for diseases in

```

```

PT which the gene is involved, or as target molecules for gene therapy.
XX
XX Claim 14; SEQ ID NO 2407; 205pp; English.
XX
XX The present invention relates to novel human secretory or membrane
XX proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
XX ADA54071). The coding sequences are useful in the gene therapy of
XX diseases caused by abnormalities of the proteins, e.g. cancer,
XX inflammatory diseases, osteoporosis or neurological disease.
XX
XX Sequence 190 AA;
XX
XX Query Match 89.5%; Score 34; DB 6; Length 190;
XX Best Local Similarity 66.7%; Pred. No. 2.5e+02;
XX Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 WXXWHF 6
XX | | | |
XX 82 WGWKWHF 87
XX
XX Db
XX
XX RESULT 41
XX AAG05810
XX ID AAG05810 standard; protein; 306 AA.
XX
XX AC AAG05810;
XX
XX DT 17-OCT-2000 (first entry)
XX
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 2354.
XX
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX OS Arabidopsis thaliana.
XX
XX PN EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX PR 25-FEB-1999; 99US-0121825P.
XX
XX PR 05-MAR-1999; 99US-0123180P.
XX
XX PR 09-MAR-1999; 99US-0123548P.
XX
XX PR 23-MAR-1999; 99US-0125788P.
XX
XX PR 25-MAR-1999; 99US-0126264P.
XX
XX PR 29-MAR-1999; 99US-0126785P.
XX
XX PR 01-APR-1999; 99US-0127462P.
XX
XX PR 06-APR-1999; 99US-0128234P.
XX
XX PR 08-APR-1999; 99US-0128714P.
XX
XX PR 16-APR-1999; 99US-0129845P.
XX
XX PR 19-APR-1999; 99US-0130077P.
XX
XX PR 21-APR-1999; 99US-0130449P.
XX
XX PR 23-APR-1999; 99US-0130510P.
XX
XX PR 28-APR-1999; 99US-0130891P.
XX
XX PR 30-APR-1999; 99US-0132048P.
XX
XX PR 30-APR-1999; 99US-0132407P.
XX
XX PR 04-MAY-1999; 99US-0132484P.
XX
XX PR 05-MAY-1999; 99US-0132485P.
XX
XX PR 06-MAY-1999; 99US-0132486P.
XX
XX PR 06-MAY-1999; 99US-0132487P.
XX
XX PR 07-MAY-1999; 99US-0132863P.
XX
XX PR 11-MAY-1999; 99US-0134256P.
XX
XX PR 14-MAY-1999; 99US-0134218P.
XX
XX PR 14-MAY-1999; 99US-0134219P.
XX
XX PR 14-MAY-1999; 99US-0134221P.
XX
XX PR 14-MAY-1999; 99US-0134370P.
XX
XX PR 18-MAY-1999; 99US-0134768P.
XX
XX PR 19-MAY-1999; 99US-0134941P.
XX
XX PR 20-MAY-1999; 99US-0135124P.

```

PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140891P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 18-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161353P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.


```
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      89.58; Score 34; DB 3; Length 306;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 WXXWHF 6
      | | | |
Db      115 WVNWHF 120

RESULT 42
AAG47194
ID AAG47194 standard; protein; 306 AA.
XX
AC AAG47194;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 59455.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01233180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 18-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
```

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161520P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match

89.5%;

Score 34; DB 3; Length 306;

Best Local Similarity

66.7%;

Pred. No. 3.8e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
Db 115 WYNWHF 120
RESULT 43
AAG05809
ID AAG05809 standard; protein; 338 AA.
XX
AC AAG05809;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2353.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130443P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155113P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 89.5%; Score 34; DB 3; Length 338;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6

Db 147 WVNWHF 152

```

RESULT 44
AAG47193
AC AAG47193 standard; protein; 338 AA.
XX
AC AAG47193;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 59454.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 18-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.

```

```
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151304P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154032P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161408P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 89.5%; Score 34; DB 3; Length 338;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 147 WYGNWHF 152

RESULT 45
ABJ25408
ID ABJ25408 standard; protein; 402 AA.
XX
XX AC ABJ25408;
XX
XX DT 16-APR-2003 (first entry)
```

```
XX Aspergillus fumigatus essential gene protein #66.
DE Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX cancer; contamination; biofilm; antibody; immune response.
KW Aspergillus fumigatus.
KW WO200286090-A2.
XX 31-OCT-2002.
XX 23-APR-2002; 2002WO-US013142.
XX 23-APR-2001; 2001US-0285697P.
PR 27-APR-2001; 2001US-0287066P.
PR 05-JUN-2001; 2001US-0295890P.
PR 09-JUL-2001; 2001US-0303899P.
PR 31-AUG-2001; 2001US-0316362P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.
XX
XX New purified or isolated nucleic acids of essential genes of Aspergillus
PT fumigatus, useful for treating or preventing infections by A. fumigatus,
PT or for treating a non-infectious disease in a subject e.g. cancer.
XX
XX Disclosure; Page; 175pp; English.
```

The invention relates to novel purified or isolated nucleic acids of essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as *A. fumigatus*, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by *A. fumigatus*, or to prevent or inhibit formation on a surface of a biofilm comprising *A. fumigatus*. The polynucleotides are useful for expressing recombinant protein for characterization, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of *A. fumigatus* to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, for raising anti-protein antibodies, as an antigen to raise anti-DNA antibodies or to elicit another immune response, and for identifying polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction. The polypeptides may be used to raise antibodies or to elicit immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case of virulence factors. This sequence represents a protein of one of the essential genes of *Aspergillus fumigatus* of the invention

```
XX Sequence 402 AA;
SQ
Query Match 89.5%; Score 34; DB 6; Length 402;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 112 WYGNWHF 117

RESULT 46
AAR60501
ID AAR60501 standard; protein; 418 AA.
```

XX AAR60501;
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 28-MAR-1995 (first entry)
 XX
 XX Linoleic-acid-desaturase.
 XX
 XX Linoleic-acid-desaturase; oleic-desaturase; transgenic plant;
 KW crop improvement; linolenic acid.
 KW
 XX Unidentified.
 OS
 XX WO9418337-A1.
 PN
 XX
 XX 18-AUG-1994.
 PD
 XX
 XX 04-FEB-1994; 94WO-US001321.
 PF
 XX 05-FEB-1993; 93US-00014431.
 PR
 XX 22-NOV-1993; 93US-00156551.
 XX
 XX (MONS) MONSANTO CO.
 PA (UNMS) UNIV MICHIGAN STATE.
 PA
 XX Gibson SI, Kishore GM, Ruff TG, Somerville CR, Arondel VJA;
 PI WPI; 1994-279758/34.
 XX N-PSDB; AAQ71242.
 DR
 XX Genetically transformed plants with altered linolenic acid content -
 PT contg recombinant, double-stranded DNA encoding linolenic acid
 PT desaturase, or the antisense of the coding sequence.
 PT
 XX Disclosure; Page 100-102; 144pp; English.
 PS
 XX The isolation of cDNA encoding linoleic-acid-desaturase and oleic-
 CC desaturase is described. A sequence of the invention is given in
 CC AAQ71242, and its encoded protein sequence in AAR60501. (Updated on 25-
 CC MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
 CC field.)
 XX
 XX Sequence 418 AA;
 SQ
 Query Match 89.5%; Score 34; DB 2; Length 418;
 Best Local Similarity 66.7%; Pred. No. 5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 Db 257 WYVWHF 262
 |||||
 RESULT 47
 ID ABB49593
 XX ABB49593 standard; protein; 423 AA.
 AC
 XX ABB49593;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Listeria monocytogenes protein #2297.
 XX
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 XX
 XX WO200177335-A2.
 PN
 XX 18-OCT-2001.
 PD
 XX
 XX
 PF 11-APR-2001; 2001WO-FR001118.
 XX
 PR 11-APR-2000; 2000FR-00004629.
 XX
 PA (INSP) INST PASTEUR.
 XX
 XX Bussieres C, Frangeul L, Couve E, Rusniok C, Faihi H, Deboux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX
 XX WPI; 2002-010914/01.
 DR
 XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 XX and prevention of Listeria and related bacterial infections, and related
 PT polypeptides.
 PT
 XX Claim 6; SEQ ID NO 2298; 192pp; French.
 PS
 XX The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 423 AA;
 SQ
 Query Match 89.5%; Score 34; DB 5; Length 423;
 Best Local Similarity 66.7%; Pred. No. 5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 Db 126 WYVWHF 131
 |||||
 RESULT 48
 ID ABU32991
 XX ABU32991 standard; protein; 423 AA.
 XX
 AC ABU32991;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #18518.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Listeria monocytogenes.
 OS
 XX WO200277183-A2.
 XX
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002WO-US0009107.
 PF
 XX
 XX

XX 03-OCT-2002.
 XX PD
 XX PF
 XX XX
 XX 21-MAR-2002; 2002WO-US009107.
 XX XX
 XX 21-MAR-2001; 2001US-00815242.
 PR PR
 PR 06-SEP-2001; 2001US-00948993.
 PR PR
 PR 23-OCT-2001; 2001US-0342923P.
 PR PR
 PR 08-FEB-2002; 2002US-00072851.
 PR PR
 PR 06-MAR-2002; 2002US-0362699P.
 XX XX
 PA (ELIT-) ELITRA PHARM INC.
 XX XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI XX
 XX WPI; 2003-029926/02.
 DR DR
 DR N-PSDB; ACA29487.
 XX XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX XX
 PS Claim 25; SEQ ID NO 53541; 1766pp; English.
 XX XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 424 AA;

Query Match 89.5%; Score 34; DB 6; Length 424;
 Best Local Similarity 66.7%; Pred. No. 5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 DB 126 WNYWHF 131

Search completed: June 10, 2004, 10:48:27
 Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:43:44 ; Search time 11.6667 Seconds
(without alignments)
49.470 Million cell updates/sec

Title: US-09-912-414-11
Perfect score: 38
Sequence: 1 WXXWHF 6

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : PIR 78:*

1: Pirl: *
2: Pirl: *
3: Pirl: *
4: Pirl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	35	92.1	241	2 A71020	hypothetical prote
2	35	92.1	358	1 S43876	site-specific DNA
3	35	92.1	358	2 A87296	modification methy
4	35	92.1	386	2 A97456	adenine-specific m
5	35	92.1	386	2 A02674	adenine DNA methyl
6	35	92.1	394	2 D90105	putative SAR DNA-b
7	35	92.1	403	2 AF3432	site-specific DNA-
8	35	92.1	497	2 T40586	nucleolar protein
9	35	92.1	522	2 D96602	nucleolar protein
10	34	89.5	230	2 D70847	hypothetical prote
11	34	89.5	423	2 AH1407	PTS system galacti
12	34	89.5	423	2 AH1783	PTS system galacti
13	34	89.5	424	2 JC5891	omega-6 desaturase
14	34	89.5	424	2 T07742	hypothetical prote
15	34	89.5	439	2 T01807	probable omega-6 d
16	34	89.5	443	2 T08136	hypothetical prote
17	34	89.5	447	2 T17299	hypothetical prote
18	34	89.5	448	2 D85362	hypothetical prote
19	34	89.5	452	2 H91172	probable PTS syste
20	34	89.5	452	2 AH0964	hypothetical prote
21	34	89.5	462	2 H86018	hypothetical prote
22	34	89.5	472	2 AG2417	hypothetical prote
23	34	89.5	475	2 B84171	hypothetical prote
24	34	89.5	487	2 T32941	hypothetical prote
25	34	89.5	504	2 S48550	hypothetical prote
26	34	89.5	508	2 G90372	4-hydroxyphenylace
27	34	89.5	508	2 T50180	nucleolar protein
28	34	89.5	511	2 S58322	nucleolar protein
29	34	89.5	550	2 T06379	SAR DNA-binding pr

30	560	89.5	560	2 T06377	SAR DNA-binding pr
31	81	86.8	81	2 S41580	lysozyme (EC 3.2.1
32	140	86.8	140	2 S20914	lysozyme (EC 3.2.1
33	140	86.8	140	2 S41573	lysozyme (EC 3.2.1
34	140	86.8	140	2 S41574	lysozyme (EC 3.2.1
35	140	86.8	140	2 S41577	lysozyme (EC 3.2.1
36	260	86.8	260	2 B71353	probable D,D-carbo
37	293	86.8	293	2 A83299	hypothetical prote
38	488	86.8	488	2 B87540	hypothetical prote
39	492	86.8	492	2 AH2079	alpha-amylase limp
40	549	86.8	549	2 T02790	hypothetical prote
41	640	86.8	640	2 T28631	hypothetical prote
42	718	86.8	718	2 T29448	hypothetical prote
43	779	86.8	779	2 AG1978	hypothetical prote
44	139	84.2	139	2 S41579	lysozyme (EC 3.2.1
45	244	84.2	244	2 T01761	hypothetical prote
46	372	84.2	372	2 D70753	probable oxidoredu
47	396	84.2	396	1 E64987	bicyclomycin resis
48	396	84.2	396	2 B91013	bicyclomycin resis
49	396	84.2	396	2 D85857	bicyclomycin resis
50	396	84.2	396	2 AF0785	bicyclomycin resis
51	441	84.2	441	2 A83370	probable MFS trans
52	462	84.2	462	2 I58179	SMT3 receptor subu
53	489	84.2	489	2 S41757	S-hydroxytryptamin
54	53	81.6	53	2 F89871	hypothetical prote
55	408	81.6	408	2 S76830	hypothetical prote
56	419	81.6	419	2 F83673	PTS system, galact
57	419	81.6	419	2 B89788	hypothetical prote
58	433	81.6	433	2 T48118	hypothetical prote
59	475	81.6	475	1 C65179	hypothetical prote
60	475	81.6	475	2 A86062	hypothetical 51.5
61	475	81.6	475	2 H91215	probable transport
62	480	81.6	480	2 S76206	probable transport
63	482	81.6	482	2 S49465	hypothetical prote
64	1099	81.6	1099	2 AE1065	catalase (EC 1.11.
65	1177	81.6	1177	2 AD0438	conserved hypotet
66	144	78.9	144	2 B70983	hypothetical prote
67	161	78.9	161	2 B82632	hypothetical prote
68	189	78.9	189	2 H82683	hypothetical prote
69	235	78.9	235	2 D75113	hypothetical prote
70	240	78.9	240	2 T37122	hypothetical prote
71	288	78.9	288	2 T37081	hypothetical prote
72	398	78.9	398	2 H72660	probable N-Acylami
73	432	78.9	432	1 FGLMGS	fibrinogen gamma c
74	591	78.9	591	2 G01586	probable protein 5
75	678	78.9	678	2 A87698	NAD(+) synthetase,
76	883	78.9	883	1 RNBPI7	DNA-directed RNA p
77	3341	78.9	3341	1 A42996	genome polyprotein
78	85	76.3	85	2 S31018	gene 73 protein -
79	99	76.3	99	2 C46518	Ig L1 chain V regi
80	106	76.3	106	2 C82715	conserved hypotet
81	108	76.3	108	2 F72507	hypothetical prote
82	110	76.3	110	2 B81147	hypothetical prote
83	110	76.3	110	2 B81876	hypothetical prote
84	112	76.3	112	2 S12155	hypothetical prote
85	113	76.3	113	2 AD0227	vpv protein - huma
86	114	76.3	114	2 G81353	probable membrane
87	126	76.3	126	2 AB1947	hypothetical prote
88	142	76.3	142	2 C75271	hypothetical prote
89	148	76.3	148	2 F86701	hypothetical prote
90	154	76.3	154	2 C91171	hypothetical prote
91	154	76.3	154	2 C86017	probable beta-hydr
92	164	76.3	164	2 E70727	hypothetical prote
93	172	76.3	172	2 D64908	hypothetical prote
94	172	76.3	172	2 B90897	hypothetical prote
95	172	76.3	172	2 E85720	hypothetical prote
96	185	76.3	185	2 T36546	hypothetical prote
97	193	76.3	193	2 A82419	conserved hypotet
98	196	76.3	196	2 T24815	hypothetical prote
99	205	76.3	205	2 B89854	conserved hypotet
100	218	76.3	218	2 T27954	hypothetical prote
101	221	76.3	221	2 T27980	hypothetical prote
102	226	76.3	226	1 F64450	hypothetical prote

103	29	76.3	254	2	T35366	probable membrane	176	29	76.3	1034	2	A95262	probable formate d
104	29	76.3	259	2	T45841	6-phosphogluconol	177	29	76.3	1164	2	T01871	RNA-directed DNA p
105	29	76.3	259	2	AD0186	probable ferric ir	178	29	76.3	1285	1	BTQPD	dermonecrotic toxi
106	29	76.3	260	2	F97272	conserved membrane	179	29	76.3	1293	2	B40025	maleless protein,
107	29	76.3	261	2	AH0673	probable secreted	180	29	76.3	1314	2	G02870	KIAA0197 protein -
108	29	76.3	263	2	H87677	conserved hypoteth	181	29	76.3	1360	2	T06699	zinc finger protei
109	29	76.3	286	2	C84538	cytochrome-c oxida	182	29	76.3	2180	2	T29764	hypothetical prote
110	29	76.3	287	2	D81295	probable Arac-fami	183	29	76.3	2638	1	A42545	genome polypotein
111	29	76.3	288	2	AG0888	cytochrome oxidase	184	29	76.3	3412	1	GNWVB	genome polypotein
112	29	76.3	292	2	G75295	hypothetical prote	185	29	76.3	3414	1	GNWNE	polypotein(NS1, N
113	29	76.3	300	2	G75295	envelope polypote	186	29	76.3	3415	2	A46105	alpha-2-macroglobu
114	29	76.3	303	2	S05550	hypothetical prote	187	29	76.3	4543	1	A53102	alpha-2-macroglobu
115	29	76.3	304	2	G27551	hypothetical prote	188	29	76.3	4544	1	S02392	alpha-2-macroglobu
116	29	76.3	312	2	JN0596	fibrinogen-related	189	29	76.3	4545	1	S25111	alpha-2-macroglobu
117	29	76.3	313	2	JC7656	LysR-type transcri	190	28	73.7	76	2	T14724	hypothetical prote
118	29	76.3	316	1	SYLBT	thymidylate syntha	191	28	73.7	79	2	A64750	hypothetical prote
119	29	76.3	323	2	S69647	hypothetical prote	192	28	73.7	86	2	T42185	hypothetical prote
120	29	76.3	360	2	T35584	probable transcrip	193	28	73.7	91	2	T11554	conserved hypoteth
121	29	76.3	384	2	T50584	probable transcrip	194	28	73.7	96	4	QOEC31	hypothetical prote
122	29	76.3	385	2	F70604	probable lipoprote	195	28	73.7	99	2	C84022	hypothetical prote
123	29	76.3	389	2	S06629	probable fadE30 pr	196	28	73.7	100	2	PQ0115	hypothetical prote
124	29	76.3	391	2	S39816	lysine acetyltrans	197	28	73.7	103	2	B70600	hypothetical prote
125	29	76.3	392	2	T49471	mucin (muc3) relat	198	28	73.7	105	2	JB0384	NADH2 dehydrogenas
126	29	76.3	395	2	B96610	hypothetical prote	199	28	73.7	105	2	H82388	hypothetical prote
127	29	76.3	397	2	D83311	conserved hypoteth	200	28	73.7	106	2	A30996	hypothetical prote
128	29	76.3	399	2	AB0155	probable drug resi	201	28	73.7	108	2	S28241	orf-X protein - si
129	29	76.3	399	2	T38388	hypothetical wd-40	202	28	73.7	111	2	AE0794	NADH2 dehydrogenas
130	29	76.3	405	2	AE2701	ferredoxin reducta	203	28	73.7	111	2	AE0794	probable membrane
131	29	76.3	405	2	F97483	redA2 protein (AJO	204	28	73.7	112	1	ASLJX2	vpu protein - huma
132	29	76.3	407	2	T19895	hypothetical prote	205	28	73.7	112	1	ASLJX	vpu protein - huma
133	29	76.3	442	2	T18917	hypothetical prote	206	28	73.7	112	1	ASLJX	vpu protein - huma
134	29	76.3	458	2	B81409	probable transmemb	207	28	73.7	112	1	ASLJST	vpu protein - simi
135	29	76.3	468	2	AD2395	two-component sens	208	28	73.7	112	1	ASLJST	vpu protein - simi
136	29	76.3	507	2	T49519	hypothetical prote	209	28	73.7	112	1	ASLJX3	vpu protein - simi
137	29	76.3	520	2	A83203	arginate o-acetyl	210	28	73.7	112	2	T11562	vpx protein - simi
138	29	76.3	550	2	S46528	probable membrane	211	28	73.7	112	2	S07990	vpx protein - simi
139	29	76.3	560	2	S64091	pectinesterase (EC	212	28	73.7	112	2	S03066	gene X protein - h
140	29	76.3	564	2	F90565	hypothetical prote	213	28	73.7	112	2	S53094	vpx protein - huma
141	29	76.3	584	2	F85813	hypothetical prote	214	28	73.7	112	2	S53094	ig heavy chain V-1
142	29	76.3	567	2	AE0382	conserved hypoteth	215	28	73.7	133	2	AF0447	probable phage-rel
143	29	76.3	567	2	AC0754	probable membrane	216	28	73.7	136	2	C95132	hypothetical prote
144	29	76.3	569	2	H64959	probable membrane	217	28	73.7	141	2	A83542	probable ring-clea
145	29	76.3	572	2	T50404	probable inorganic	218	28	73.7	149	2	D84181	hypothetical prote
146	29	76.3	572	2	S65208	ATP-dependent prot	219	28	73.7	171	2	T35339	hypothetical prote
147	29	76.3	598	2	C82194	probable membrane	220	28	73.7	171	2	D70830	hypothetical prote
148	29	76.3	609	2	S65208	potassium transpor	221	28	73.7	171	2	D70830	hypothetical prote
149	29	76.3	622	2	AG0001	flagellar hook-ass	222	28	73.7	174	2	F75606	hypothetical prote
150	29	76.3	627	2	E70122	sensory transducti	223	28	73.7	175	2	G69856	hypothetical prote
151	29	76.3	629	2	T05089	hypothetical prote	224	28	73.7	177	2	A83753	hypothetical prote
152	29	76.3	633	2	S75525	hypothetical prote	225	28	73.7	178	2	A86836	hypothetical prote
153	29	76.3	676	2	T32556	hypothetical prote	226	28	73.7	185	2	AD1026	hypothetical prote
154	29	76.3	701	2	T19605	related to SHK1 KI	227	28	73.7	185	2	T44538	hypothetical prote
155	29	76.3	718	2	T49572	probable outer mem	228	28	73.7	194	2	S33294	translocon-associa
156	29	76.3	720	2	F85572	probable outer mem	229	28	73.7	194	2	A82198	hypothetical prote
157	29	76.3	723	2	G70721	hypothetical prote	230	28	73.7	209	2	E70536	hypothetical prote
158	29	76.3	729	2	G70539	hypothetical prote	231	28	73.7	221	2	B72506	hypothetical prote
159	29	76.3	779	2	AE2402	alpha-glucosidase	232	28	73.7	232	2	E82140	secretory protein
160	29	76.3	780	2	T29680	hypothetical prote	233	28	73.7	234	2	UH0483	hypothetical prote
161	29	76.3	798	2	T23539	hypothetical prote	234	28	73.7	239	2	A12299	transcription regu
162	29	76.3	815	2	AG2278	cation-transportin	235	28	73.7	239	2	A12299	hypothetical prote
163	29	76.3	818	1	E64807	outer membrane uah	236	28	73.7	242	2	AC1646	inner membrane prot
164	29	76.3	827	2	AB3764	beta-mannosidase p	237	28	73.7	254	2	E97358	60s ribosomal prot
165	29	76.3	832	2	A97545	hypothetical prote	238	28	73.7	255	2	H90130	conserved hypoteth
166	29	76.3	832	2	H84848	phospholipase D li	239	28	73.7	255	2	A69433	trRNA guanine-N1 me
167	29	76.3	833	2	T01547	probable phospholi	240	28	73.7	257	2	E70429	toxin cytb - Bacil
168	29	76.3	931	2	T33744	hypothetical prote	241	28	73.7	259	2	S32432	hypothetical prote
169	29	76.3	974	2	ES9434	Rho GTPase activat	242	28	73.7	260	2	S76509	integral membrane
170	29	76.3	1015	1	JS0628	formate dehydrogen	243	28	73.7	260	2	T18554	EKSI protein - Yea
171	29	76.3	1015	2	C85729	hypothetical prote	244	28	73.7	262	2	S22850	glucose 1-dehydrog
172	29	76.3	1015	2	F90888	hypothetical prote	245	28	73.7	262	2	B70073	probable permease
173	29	76.3	1016	1	S40838	formate dehydrogen	246	28	73.7	265	2	AC0441	probable aliphatic
174	29	76.3	1016	2	D91231	formate dehydrogen	247	28	73.7	265	2	T01578	probable membrane
175	29	76.3	1021	2	AB0946	probable DNA-direc	248	28	73.7	266	2	AD3238	hypothetical prote

249	28	73.7	267	1	S71020	peptidoglycan-link	322	28	73.7	421	1	JC4277	ribosomal protein
250	28	73.7	267	2	C81937	competence lipopro	323	28	73.7	421	1	JC4277	probable phosphatase
251	28	73.7	267	2	H81167	competence lipopro	324	28	73.7	426	1	T09551	ribosomal protein
252	28	73.7	270	2	C86295	hypothetical prote	325	28	73.7	427	2	S55905	phosphotransferase
253	28	73.7	272	2	H70510	hypothetical prote	326	28	73.7	434	2	T47748	alpha-galactosidase
254	28	73.7	272	2	H63182	hypothetical prote	327	28	73.7	434	2	H70989	probable GABA perm
255	28	73.7	275	2	S76778	hypothetical prote	328	28	73.7	435	2	E70711	hypothetical prote
256	28	73.7	283	2	T36769	hypothetical prote	329	28	73.7	435	2	G90700	probable uracil tr
257	28	73.7	283	2	T14455	hypothetical prote	330	28	73.7	435	2	B85551	probable transport
258	28	73.7	286	2	AD1919	hypothetical prote	331	28	73.7	435	2	H84782	probable membrane
259	28	73.7	287	2	H90674	hypothetical prote	332	28	73.7	440	2	JC2532	secretin receptor
260	28	73.7	287	2	C85525	hypothetical prote	333	28	73.7	441	2	G83212	hypothetical prote
261	28	73.7	287	2	G64758	yahE protein - Esc	334	28	73.7	446	2	B89922	conserved hypotet
262	28	73.7	293	2	AG2929	hypothetical prote	335	28	73.7	449	2	S16319	secretin receptor
263	28	73.7	293	2	G98352	hypothetical prote	336	28	73.7	450	2	B82973	conserved hypotet
264	28	73.7	296	2	B82704	dipeptide ABC tran	337	28	73.7	451	1	JC6180	secretin receptor
265	28	73.7	296	2	D84240	arginine deaminase	338	28	73.7	451	2	C64976	conserved hypotet
266	28	73.7	297	2	A2596	hypothetical prote	339	28	73.7	451	2	G90990	stearyl-CoA 9-Des
267	28	73.7	298	2	A10942	major outer membra	340	28	73.7	451	2	A85836	gacC protein - Esc
268	28	73.7	300	2	S40858	probable transmembr	341	28	73.7	452	2	T28094	PTS system galacti
269	28	73.7	300	2	G96620	hypothetical prote	342	28	73.7	452	2	T39383	hypothetical prote
270	28	73.7	302	2	H86080	hypothetical prote	343	28	73.7	454	2	T13886	angio-associated m
271	28	73.7	302	2	H31233	probable transport	344	28	73.7	455	2	B71480	NADH2 dehydrogenas
272	28	73.7	302	2	A33308	probable transport	345	28	73.7	455	2	B81743	probable agx-1 hom
273	28	73.7	303	2	S60549	cobalt-zinc-cadmiu	346	28	73.7	455	2	AH0898	UMP-N-acetylglucos
274	28	73.7	307	2	G83030	envelope polypote	347	28	73.7	459	2	JH0594	PTS system, galact
275	28	73.7	310	2	A25027	hypothetical prote	348	28	73.7	460	2	JC2194	vasoactive intesti
276	28	73.7	311	2	A80275	arabinose operon r	349	28	73.7	460	2	C82964	vasoactive intesti
277	28	73.7	312	2	D86435	arabinose operon r	350	28	73.7	461	1	KWHUFA	glycosyltransferas
278	28	73.7	316	2	C47099	unknown protein [i	351	28	73.7	462	2	S10235	alpha-L-fucosidase
279	28	73.7	316	2	C64963	lysR homolog, nac	352	28	73.7	462	2	AB3536	penicillin amidase
280	28	73.7	316	2	E85823	transcription regu	353	28	73.7	466	2	T44252	transport protein
281	28	73.7	319	2	G90976	transcription regu	354	28	73.7	469	2	T50934	dioxigenase D1A1,
282	28	73.7	323	2	G83279	probable hydrolase	355	28	73.7	473	2	H97092	hypothetical prote
283	28	73.7	323	2	AC0117	hypothetical prote	356	28	73.7	476	1	A35104	brain-derived neur
284	28	73.7	327	2	A83766	hypothetical prote	357	28	73.7	476	1	I73631	brain-derived neur
285	28	73.7	329	2	AR1891	hypothetical prote	358	28	73.7	477	2	AE2313	hypothetical prote
286	28	73.7	339	2	T32291	hypothetical prote	359	28	73.7	478	2	C81103	hypothetical prote
287	28	73.7	345	2	C82270	hypothetical prote	360	28	73.7	481	1	VCB913	alginate O-acetyla
288	28	73.7	350	2	D65018	hypothetical prote	361	28	73.7	483	2	H81838	probable polysacch
289	28	73.7	350	2	D91042	Ethanolamine opero	362	28	73.7	485	2	S56359	minor coat protein
290	28	73.7	350	2	G85886	probable ARAC-type	363	28	73.7	485	2	S66108	probable proton/ol
291	28	73.7	350	2	AC0813	probable ARAC-type	364	28	73.7	485	2	H91267	probable peptide t
292	28	73.7	356	2	E75575	ethanolamine opero	365	28	73.7	486	2	AG0826	probable peptide t
293	28	73.7	357	2	C90772	probable sulfite o	366	28	73.7	495	2	JC2195	probable transmembr
294	28	73.7	357	2	G85634	hypothetical prote	367	28	73.7	495	2	D71308	vasoactive intesti
295	28	73.7	358	2	S41640	hypothetical prote	368	28	73.7	495	2	JC1308	probable alginate
296	28	73.7	363	2	T40797	ribosomal protein	369	28	73.7	496	2	S55273	amine oxidase (fla
297	28	73.7	370	2	A96741	60S ribosomal prot	370	28	73.7	499	2	B83591	probable transport
298	28	73.7	370	2	E82369	hypothetical prote	371	28	73.7	502	2	AH2413	alginate O-acetyl
299	28	73.7	374	2	AC0983	thiH protein VC006	372	28	73.7	513	2	T17416	probable alkylhali
300	28	73.7	374	2	C87393	hypothetical prote	373	28	73.7	513	2	T42050	probable transposa
301	28	73.7	384	2	S37608	hypothetical prote	374	28	73.7	527	2	G64626	alginate O-acetyla
302	28	73.7	384	2	E75295	NADH2 dehydrogenas	375	28	73.7	533	2	B67777	hypothetical prote
303	28	73.7	395	2	T19201	conserved hypotet	376	28	73.7	533	2	AG2293	hypothetical prote
304	28	73.7	386	2	T12048	hypothetical prote	377	28	73.7	543	2	S64850	probable membrane
305	28	73.7	394	2	C64185	ribosomal protein	378	28	73.7	558	2	A42463	hypothetical prote
306	28	73.7	394	2	AC1554	cell division prot	379	28	73.7	558	2	S75104	hypothetical prote
307	28	73.7	394	2	AE1196	DltB protein for D	380	28	73.7	565	2	G82443	conserved hypotet
308	28	73.7	395	2	S39659	DltB protein for D	381	28	73.7	606	2	T08180	PF20 protein, micr
309	28	73.7	396	1	R5X11A	dltB protein - Bac	382	28	73.7	614	1	S75294	ferrous iron trans
310	28	73.7	396	1	R5X11B	ribosomal protein	383	28	73.7	624	2	T33868	hypothetical prote
311	28	73.7	398	2	AB2669	conserved hypotet	384	28	73.7	630	2	JC5374	angiotensin-conver
312	28	73.7	398	2	H97450	hypothetical prote	385	28	73.7	640	2	T26820	hypothetical prote
313	28	73.7	400	2	B97020	argininosuccinate	386	28	73.7	646	2	T41545	hypothetical trans
314	28	73.7	401	2	T32167	hypothetical prote	387	28	73.7	655	2	E84649	hypothetical prote
315	28	73.7	403	2	D82076	tRNA nucleotidyltr	388	28	73.7	686	2	S70180	chemotaxis protein
316	28	73.7	404	2	F89859	DltB membrane prot	389	28	73.7	700	2	H82428	Na+/H+ antiporter
317	28	73.7	407	2	T48308	60S ribosomal prot	390	28	73.7	703	2	AF2414	iron(III) ABC tran
318	28	73.7	407	2	D86782	peptidoglycan bios	391	28	73.7	705	2	C75118	subtilase family p
319	28	73.7	407	2	S44171	phenoxycarbonate di	392	28	73.7	707	2	S77094	glycogen operon pr
320	28	73.7	413	2	D95254	dltB protein limpo	393	28	73.7	709	2	T16584	hypothetical prote
321	28	73.7	414	2	B98119	hypothetical prote	394	28	73.7	709	2	T16584	hypothetical prote

C:Accession: S43876

R; Zweiger, G.; Marczynski, G.; Shapiro, L.

J. Mol. Biol. 235, 472-485, 1994

A:Title: A Caulobacter DNA methyltransferase that functions only in the predivisional cell

A:Reference number: S43876; PMID:8289276

A:Accession: S43876

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-358 <ZWE>

A:Cross-references: EMBL:U01032; NID:G393011; PIDN:AA18913.1; PID:G393012

C:Superfamily: type II site-specific DNA-methyltransferase

C:Keywords: methyltransferase; restriction modification system

Query Match 92.1%; Score 35; DB 1; Length 358;

Best Local Similarity 66.7%; Pred. No. 68;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 332 WTYWHF 337

RESULT 3

A87296

modification methylase CcrMI [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 22-Jun-2003

C:Accession: A87296

R; Nicrman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

n. J.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; PMID:21173698; PMID:11259647

A:Accession: A87296

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-358 <STO>

A:Cross-references: GB:AE005673; NID:G13421535; PIDN:AAK2365.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0378

C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 92.1%; Score 35; DB 2; Length 358;

Best Local Similarity 66.7%; Pred. No. 68;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 332 WTYWHF 337

RESULT 4

A97456

adenine-specific methyltransferase SMEI (modification methylase SMEI) [imported] - Agrob

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Jun-2003

C:Accession: A97456

R; Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; PMID:21608551; PMID:11743194

A:Accession: A97456

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-386 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK86602.1; PID:G15155772; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C1453

A:Map position: circular chromosome

C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 92.1%; Score 35; DB 2; Length 386;

Best Local Similarity 66.7%; Pred. No. 72;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 357 WTFWHF 362

RESULT 5

AD2674

adenine DNA methyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupo

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 22-Jun-2003

C:Accession: AD2674

R; Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCle

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: AD2674

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-386 <KUR>

A:Cross-references: GB:AE008688; PIDN:AA141810.1; PID:G17739166; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

A:Genetics:

A:Gene: Atu0794

A:Map position: circular chromosome

C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 92.1%; Score 35; DB 2; Length 386;

Best Local Similarity 66.7%; Pred. No. 72;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 357 WTFWHF 362

RESULT 6

D90105

putative SAR DNA-binding protein-1 [imported] - Guillardia theta nucleomorph

C:Species: nucleomorph Guillardia theta

A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001

C:Accession: D90105

R; Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; R

Nature 410, 1091-1096, 2001

A:Title: The highly reduced genome of an enslaved algal nucleus.

A:Reference number: A99082; PMID:11323671; PMID:11323671

A:Accession: D90105

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-394 <DOU>

A:Cross-references: GB:AJ010592; NID:G12580672; PIDN:CAC26989.1; GSPDB:GN00151

C:Genetics:

A:Map position: 2

A:Genome: nucleomorph

C:Keywords: nucleomorph

Query Match 92.1%; Score 35; DB 2; Length 394;

Best Local Similarity 66.7%; Pred. No. 74;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 169 WYSWHF 174

```

RESULT 7
AF3432
site-specific DNA-methyltransferase (adenine-specific) (EC 2.1.1.72) [imported] - Brucella
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 22-Jun-2003
C:Accession: AF3432
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AB2252; PMID:11756688
A:Accession: AF3432
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-403 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52625.1; PID:g17983446; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1444
A:Map position: 1
C:Superfamily: type II site-specific DNA-methyltransferase
C:Keywords: methyltransferase; S-adenosylmethionine

Query Match          92.1%; Score 35; DB 2; Length 403;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 373 WTFWHF 378

RESULT 8
T40586
nucleolar protein involved in pre-rRNA processing - fission yeast (Schizosaccharomyces P
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T40586
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 1999
A:Reference number: Z21938
A:Accession: T40586
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-497 <SEE>
A:Cross-references: EMBL:AL035216; PIDN:CAA22814.1; GSPDB:GN00067; SPDB:SPBC646.10C
A:Experimental source: strain 972h-; cosmid c646
C:Genetics:
A:Gene: SPDB:SPBC646.10C
A:Map position: 2
C:Superfamily: garden pea SAR DNA-binding protein

Query Match          92.1%; Score 35; DB 2; Length 497;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 194 WTSWHF 199

RESULT 9
D96602
nucleolar protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana [mouse-ear cross]
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D96602
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

```

```

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96602
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <STO>
A:Cross-references: GB:AE005173; NID:g6056371; PIDN:AAF02835.1; GSPDB:GN00141
C:Genetics:
A:Gene: T6H22.10
A:Map position: 1
C:Superfamily: garden pea SAR DNA-binding protein

Query Match          92.1%; Score 35; DB 2; Length 522;
Best Local Similarity 66.7%; Pred. No. 94;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 195 WYSWHF 200

RESULT 10
D70847
hypothetical protein RV0059 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: D70847
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70847
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-230 <COL>
A:Cross-references: GB:AL021428; GB:AL123456; NID:g3261514; PIDN:CAA16240.1; PID:g280
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0059
C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV0059

Query Match          89.5%; Score 34; DB 2; Length 230;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 26 WIVWHF 31

RESULT 11
AH1407
PTS system galactitol-specific enzyme IIC component homolog lmo2665 [imported] - List
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AH1407
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioeci
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Kars, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1407
A:Status: preliminary
A:Molecule type: DNA

```

A;Residues: 1-423 <GLA>
 A;Cross-references: GB:NC_003210; PIDN:CAD00878.1; PID:G16412165; GSPDB:GN00177
 A;Experimental source: strain EGD-e
 C:Genetics:
 A;Gene: lmo2665
 C;Superfamily: phosphotransferase enzyme II galactitol-specific

Query Match 89.5%; Score 34; DB 2; Length 423;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 |
 |
 |
 Db 126 WNYWHF 131

RESULT 12
 AH1783
 PTS system galactitol-specific enzyme IIC component homolog lin2814 [imported] - Listeria
 C:Species: Listeria innocua
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C;Accession: AH1783
 R;Glaser, P.; Frangeul, L.; Ruchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournan, A.; Meok, C.; Schluter, T.; Simoes, N.; Rierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AH1783
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-423 <GLA>
 A;Cross-references: GB:AL592022; PIDN:CAC98040.1; PID:G16415350; GSPDB:GN00178
 A;Experimental source: strain Clp11262
 C:Genetics:
 A;Gene: lin2814
 C;Superfamily: phosphotransferase enzyme II galactitol-specific

Query Match 89.5%; Score 34; DB 2; Length 423;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 |
 |
 |
 Db 126 WNYWHF 131

RESULT 13
 JC5891
 omega 6 desaturase (EC 1.14.99.-) precursor, chloroplast - Chlamydomonas reinhardtii
 C:Species: Chlamydomonas reinhardtii
 C;Date: 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 20-Jun-2000
 C;Accession: JC5891
 R;Sato, N.; Fujiwara, S.; Kawaguchi, A.; Tsuzuki, M.
 J. Biochem. 122, 1224-1232, 1997
 A;Title: Cloning of a gene for chloroplast omega 6 desaturase of a green alga, Chlamydomonas reinhardtii.
 A;Reference number: JC5891; MUID:98158334; PMID:9498569
 A;Accession: JC5891
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-424 <SAT>
 A;Cross-references: DBJ:AB007640; NID:G2696716; PIDN:BAA23881.1; PID:G2696717
 C;Comment: This enzyme catalyzes the desaturation of monoenoic to dienoic acids.
 C:Genetics:
 A;Gene: des6
 C;Superfamily: omega-3 fatty acid desaturase

C;Keywords: chloroplast; oxidoreductase
 F1-40/Domain: transit peptide (chloroplast) #status predicted <TPS>

Query Match 89.5%; Score 34; DB 2; Length 424;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 |
 |
 |
 Db 234 WLVWHF 239
 RESULT 14
 T07742
 omega-6 desaturase, chloroplast - soybean
 C:Species: Glycine max (soybean)
 C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 08-Oct-1999
 C;Accession: T07742
 R;Hitz, W.D.; Carlson, T.J.; Booth, J.R.; Kinney, A.J.; Stecca, K.L.; Yadav, N.S.
 Plant Physiol. 105, 635-641, 1994
 A;Title: Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA and its expression in Arabidopsis.
 A;Reference number: Z16109; MUID:94345008; PMID:8066133
 A;Accession: T07742
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-424 <HIT>
 A;Cross-references: EMBL:L29215; NID:G459961; PIDN:AAA50158.1; PID:G459962
 A;Experimental source: seed
 C:Genetics:
 A;Genome: nuclear
 C;Superfamily: omega-3 fatty acid desaturase
 C;Keywords: chloroplast

Query Match 89.5%; Score 34; DB 2; Length 424;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 |
 |
 |
 Db 251 WLVWHF 256

RESULT 15
 T01807
 hypothetical protein A.TM021B04.12 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 02-Jun-2000
 C;Accession: T01807
 R;pante, M.; Wamsley, P.; Gibson, A.
 submitted to the EMBL Data Library, June 1997
 A;Description: The sequence of A. thaliana TM021B04.
 A;Reference number: Z14440
 A;Accession: T01807
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-439 <DAN>
 A;Cross-references: EMBL:AF007271; NID:G2191181; PID:G2191188; GSPDB:GN00063; ATSP:A_1
 C:Genetics:
 A;Gene: ATSP:A.TM021B04.12
 A;Map position: 5
 A;Introns: 26/3; 40/3; 100/3; 155/3; 222/3; 388/3
 C;Superfamily: garden pea SAR DNA-binding protein

Query Match 89.5%; Score 34; DB 2; Length 439;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 |
 |
 |
 Db 184 WFGWHF 189

RESULT 16
 T08136
 probable omega-6 desaturase (EC 1.14.99.-) precursor, chloroplast - rape
 C:Species: Brassica napus (rape)
 C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 08-Oct-1999
 C;Accession: T08136

R;Hitz, W.D.; Carlson, T.J.; Booth, J.R.; Kinney, A.J.; Stecca, K.L.; Yadav, N.S.
Plant Physiol. 105, 635-641, 1994
A;Title: Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA and its ex
A;Reference number: Z16109; MUID:94345008; PMID:8066133
A;Accession: T08136
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-443 <HIT>
A;Cross-references: EMBL:L29214; NID:9457630; PIDN:AA50157.1; PID:9457631
A;Experimental source: seed
C;Genetics:
A;Genome: nuclear
C;Superfamily: Omega-3 fatty acid desaturase
C;Keywords: chloroplast; fatty acid metabolism; oxidoreductase

Query Match 89.5%; Score 34; DB 2; Length 443;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 252 WYWHF 257

RESULT 17

T17239

hypothetical protein DKFp564H2171.1 - human (fragment)

C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: T17299
R;Boecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18724
A;Accession: T17299
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-447 <BLO>
A;Cross-references: EMBL:AL117554
A;Experimental source: fetal brain; clone DKFp564H2171
C;Genetics:
A;Note: DKFp564H2171.1
C;Superfamily: garden pea SAR DNA-binding protein

Query Match 89.5%; Score 34; DB 2; Length 447;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 186 WYWHF 191

RESULT 18

D85362

hypothetical protein AT4g30950 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C;Accession: D85362
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: D85362
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-448 <STO>
A;Cross-references: GB:NC_001268; NID:97269997; PIDN:CAB79813.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g30950
A;Map position: 4
C;Superfamily: omega-3 fatty acid desaturase

Query Match 89.5%; Score 34; DB 2; Length 448;

QY 1 WXXWHF 6
DB 130 WYWHF 135

Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 257 WYWHF 262

RESULT 19

H91172

probable phosphotransferase system enzyme IIC [imported] - Escherichia coli (strain O
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: H91172
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H91172
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-452 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA937775.1; PID:g13363826; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: ECs4352
C;Superfamily: phosphotransferase enzyme II galactitol-specific

Query Match 89.5%; Score 34; DB 2; Length 452;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 130 WYWHF 135

RESULT 20

AH0964

probable PTS system IIC component STV4000 [imported] - Salmonella enterica subsp. ent
C;Species: Salmonella enterica subsp. enterica serovar typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH0964
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, J.
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0964
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-452 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD03212.1; PID:g16504843; GSPDB:GN00176
C;Genetics:
A;Gene: STY4000
C;Superfamily: phosphotransferase enzyme II galactitol-specific

Query Match 89.5%; Score 34; DB 2; Length 452;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 130 WYWHF 135

RESULT 21

H86018

hypothetical protein Z4877 [imported] - Escherichia coli (strain O157:H7, substrain E)

C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C;Accession: H86018
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamou, K.; Apodaca,
 Nature 409, 523-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: H86018
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-462 <STO>
 A;Cross-references: GB:AB005174; NID:G12518168; PIDN:AAG58612.1; GSPDB:GN00145; UWGP:Z48
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: Z4877
 C;Superfamily: phosphotransferase enzyme II galactitol-specific

Query Match 89.5%; Score 34; DB 2; Length 462;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 Db 140 WNYWHF 145

RESULT 22
 AG2417
 hypothetical protein all4895 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: AG2417
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuri, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AG2417
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-472 <KUR>
 A;Cross-references: GB:BA000019; PIDN:BAH76594.1; PID:G17134033; GSPDB:GN00179
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: all4895

Query Match 89.5%; Score 34; DB 2; Length 472;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 Db 285 WYQWHF 290

RESULT 23
 B84171
 hypothetical protein rmes [imported] - Halobacterium sp. NRC-1
 C;Species: Halobacterium sp. NRC-1
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: B84171
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
 i; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
 A;Title: Genome sequence of Halobacterium species NRC-1.
 A;Reference number: A84160; MUID:20504483; PMID:11016950
 A;Accession: B84171
 A;Status: preliminary
 A;Molecule type: DNA

A;Residues: 1-475 <STO>
 A;Cross-references: GB:AB004437; NID:G10579756; PIDN:AAG18734.1; GSPDB:GN00138
 C;Genetics: rmes
 A;Gene: rmes

Query Match 89.5%; Score 34; DB 2; Length 475;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 Db 81 WDGWHF 86

RESULT 24

T32941

hypothetical protein W01B11.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Jun-2000

C;Accession: T32941

R;Bradshaw, H.; Graves, T.; Blair, T.

submitted to the EMBL Data Library, January 1998

A;Description: The sequence of C. elegans cosmid W01B11.

A;Reference number: Z21250

A;Accession: T32941

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-487 <BRA>

A;Cross-references: EMBL:AF043704; PIDN:AAB97597.1; GSPDB:GN00019; CESP:W01B11.3

A;Experimental source: strain Bristol N2; clone W01B11

C;Genetics:

A;Gene: CESP:W01B11.3

A;Map position: 1

A;Introns: 41/2; 204/3; 266/2; 355/3

C;Superfamily: garden pea SAR DNA-binding protein

Query Match

89.5%; Score 34; DB 2; Length 487;

Best Local Similarity 66.7%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 184 WYQWHF 189

RESULT 25

S48550

hypothetical protein YLR197w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein L8167.9

C;Species: Saccharomyces cerevisiae

C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 02-Jun-2000

C;Accession: S48550

R;Fauley, A.

submitted to the EMBL Data Library, September 1994

A;Description: The sequence of S. cerevisiae cosmid 8167.

A;Reference number: S48545

A;Accession: S48550

A;Molecule type: DNA

A;Residues: 1-504 <PAU>

A;Cross-references: EMBL:U14913; NID:G544497; PIDN:AAB67431.1; PID:G544506; MIPS:YLR1.

C;Genetics:

A;Gene: SGD:SIK1

A;Cross-references: SGD:S0004187; MIPS:YLR197w

A;Map position: 12R

C;Superfamily: garden pea SAR DNA-binding protein

Query Match

89.5%; Score 34; DB 2; Length 504;

Best Local Similarity 66.7%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 198 WYQWHF 203

RESULT 26

G90372

4-hydroxyphenylacetate-3-hydroxylase (hpaA) [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 01-Mar-2002

C:Accession: G90372

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.

arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: G90372

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-508 <KUR>

A:Cross-references: GB:AE006641; NID:gl3815338; PIDN:AAK42238.1; GSPDB:GN00155

C:Genetics:

A:Gene: hpaA

C:Superfamily: Escherichia coli 4-hydroxyphenylacetate 3-monoxygenase large chain

Query Match

Best Local Similarity 89.5%; Score 34; DB 2; Length 508;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

DB 310 WFNWHF 315

RESULT 27

T50180

nucleolar protein NOP5-like protein [imported] - fission yeast (Schizosaccharomyces pombe

C:Species: Schizosaccharomyces pombe

C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000

C:Accession: T50180

R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 2000

A:Reference number: Z25044

A:Accession: T50180

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-508 <BAD>

A:Cross-references: EMBL:AL139854; PIDN:CAR72231.1; GSPDB:GN00066; SPDB:SPAC23G3.06

A:Experimental source: strain 972h(-); cosmid c23G3

C:Genetics:

A:Gene: SPDB:SPAC23G3.06

A:Map position: 1

A:introns: 6/1; 40/2

C:Superfamily: garden pea SAR DNA-binding protein

Query Match

Best Local Similarity 89.5%; Score 34; DB 2; Length 508;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

DB 187 WYGWHP 192

RESULT 28

S58322

nucleolar protein NOP5 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein O6108; protein YOR310c

C:Species: Saccharomyces cerevisiae

C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000

C:Accession: S58322; S67216; S71990

R:Pearson, B.M.; Hernandez, Y.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.

submitted to the EMBL Data Library, August 1995

A:Reference number: S58318

A:Accession: S58322

A:Molecule type: DNA

A:Residues: 1-511 <PEA>

A:Cross-references: EMBL:X90565; NID:g940836; PIDN:CAA62165.1; PID:g940841

R:Pearson, B.M.; Hernandez, Y.; Kalogeropoulos, A.; Schweizer, M.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67213

A:Accession: S67216

A:Molecule type: DNA

A:Residues: 1-511 <PEW>

A:Cross-references: EMBL:Z75217; NID:gl420680; PIDN:CAA99630.1; PID:gl420682; MIPS:Y01

A:Experimental source: strain S288C

R:Pearson, B.M.; Hernandez, Y.; Payne, J.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.

Yeast 12, 1021-1031, 1996

A:Title: Sequencing of a 35.71 kb DNA segment on the right arm of yeast chromosome XV

A:Reference number: S71986; MUID:97051589; PMID:8896266

A:Accession: S71990

A:Molecule type: DNA

A:Residues: 1-511 <PEF>

A:Cross-references: EMBL:X90565; NID:g940836; PIDN:CAA62165.1; PID:g940841

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995

C:Genetics:

A:Gene: SGD:NOP5

A:Cross-references: SGD:S0005837; MIPS:YOR310c

A:Map position: 15R

C:Function:

A:Description: involved in the synthesis of the 40S ribosomal subunit

C:Superfamily: garden pea SAR DNA-binding protein

C:Keywords: nucleus

Query Match

Best Local Similarity 89.5%; Score 34; DB 2; Length 511;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

DB 187 WYGWHP 192

RESULT 29

T06379

SAR DNA-binding protein 2 - garden pea

C:Species: Pisum sativum (garden pea)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 02-Jun-2000

C:Accession: T06379

R:Hutton, D.; Gray, J.C.

submitted to the EMBL Data Library, April 1998

A:Description: cDNA encoding a pea SAR DNA-binding protein that shows homology to nuc

A:Reference number: Z15637

A:Accession: T06379

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-550 <HAT>

A:Cross-references: EMBL:AF061963; NID:g3132697; PIDN:AAC16331.1; PID:g3132698

C:Genetics:

A:Gene: SAREP-2

C:Superfamily: garden pea SAR DNA-binding protein

Query Match

Best Local Similarity 89.5%; Score 34; DB 2; Length 550;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

DB 184 WYGWHP 189

RESULT 30

T06377

SAR DNA-binding protein-1 - garden pea

C:Species: Pisum sativum (garden pea)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 02-Jun-2000

C:Accession: T06377

R;Batton, D.; Gray, J. C.
 submitted to the EMBL Data Library, April 1998
 A;Description: cDNA encoding a pea SAR DNA-binding protein that shows homology to nucle
 A;Reference number: Z15637
 A;Accession: T06377
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-560 <MAT>
 A;Cross-references: EMBL:AF061962; NID:g3132695; PIDN:AAC16330.1; PID:g3132696
 C;Genetics:
 A;Gene: SARBP-1
 C;Superfamily: garden pea SAR DNA-binding protein

Query Match 89.5%; Score 34; DB 2; Length 560;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 184 WYGVHF 189

RESULT 31

S41580
 Lysozyme (EC 3.2.1.17) X - fruit fly (Drosophila melanogaster) (fragment)
 C;Species: Drosophila melanogaster
 C;Date: 25-Dec-1994 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
 C;Accession: S41580; S32650
 R;Daffre, S.; Kilsten, P.; Samakovlis, C.; Hultmark, D.
 Mol. Gen. Genet. 242, 152-162, 1994
 A;Title: The lysozyme locus in Drosophila melanogaster: an expanded gene family adapted
 A;Reference number: S41573; MUID:94211204; PMID:8159165
 A;Accession: S41580
 A;Molecule type: mRNA
 A;Residues: 1-81 <DAF>
 A;Cross-references: EMBL:Z22224; NID:g288922; PIDN:CAA80226.1; PID:g288923
 C;Genetics:
 A;Gene: lysp
 A;Cross-references: FlyBase:FBgn0004431
 A;Map position: 3
 C;Superfamily: lysozyme c
 C;Keywords: bacteriolytic enzyme; glycosidase; hydrolase; polysaccharide degradation

Query Match 86.8%; Score 33; DB 2; Length 81;
 Best Local Similarity 50.0%; Pred. No. 38;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 63 WSTWHY 68

RESULT 32

S20914
 Lysozyme (EC 3.2.1.17) D precursor - fruit fly (Drosophila melanogaster)
 C;Species: Drosophila melanogaster
 C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999
 C;Accession: S20914; S41576
 R;Kylsten, P.; Kimbrell, D.A.; Daffre, S.; Samakovlis, C.; Hultmark, D.
 Mol. Gen. Genet. 232, 335-343, 1992
 A;Title: The lysozyme locus in Drosophila melanogaster: different genes are expressed in
 A;Reference number: S20914; MUID:92269751; PMID:1588905
 A;Accession: S20914
 A;Molecule type: DNA
 A;Residues: 1-140 <KYL>
 A;Cross-references: EMBL:X58382; NID:g8197; PIDN:CAA41272.1; PID:g8198
 R;Daffre, S.; Kilsten, P.; Samakovlis, C.; Hultmark, D.
 Mol. Gen. Genet. 242, 152-162, 1994
 A;Title: The lysozyme locus in Drosophila melanogaster: an expanded gene family adapted
 A;Reference number: S41573; MUID:94211204; PMID:8159165
 A;Accession: S41576
 A;Molecule type: mRNA
 A;Residues: 1-140 <DAF>

C;Genetics:
 A;Gene: lysd
 A;Cross-references: FlyBase:FBgn0004427
 A;Map position: 3
 C;Superfamily: lysozyme c
 C;Keywords: bacteriolytic enzyme; glycosidase; hydrolase; polysaccharide degradation
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-140/Product: lysozyme D #status predicted <MAT>

Query Match 86.8%; Score 33; DB 2; Length 140;
 Best Local Similarity 50.0%; Pred. No. 62;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 123 WSTWHY 128

RESULT 33

S41573
 Lysozyme (EC 3.2.1.17) precursor - fruit fly (Drosophila melanogaster)
 C;Species: Drosophila melanogaster
 C;Date: 25-Dec-1994 #sequence_revision 23-Feb-1996 #text_change 22-Jun-1999
 C;Accession: S41573; S41575; S32651; S32726
 R;Daffre, S.; Kilsten, P.; Samakovlis, C.; Hultmark, D.
 Mol. Gen. Genet. 242, 152-162, 1994
 A;Title: The lysozyme locus in Drosophila melanogaster: an expanded gene family adapte
 A;Reference number: S41573; MUID:94211204; PMID:8159165
 A;Accession: S41573
 A;Molecule type: mRNA
 A;Residues: 1-140 <DAF>
 A;Cross-references: EMBL:Z22223; NID:g288920; PIDN:CAA80225.1; PID:g288921
 C;Genetics: LYSB
 A;Accession: S41575
 A;Molecule type: mRNA
 A;Residues: 1-140 <DAA>
 A;Cross-references: EMBL:Z22226; NID:g296038; PIDN:CAA80228.1; PID:g296039
 C;Genetics: LYSC
 C;Genetics: <LYSA>
 A;Gene: LysA
 A;Map position: 3
 C;Genetics: <LYSC>
 A;Gene: LysC
 A;Map position: 3
 C;Superfamily: lysozyme c
 C;Keywords: bacteriolytic enzyme; glycosidase; hydrolase; polysaccharide degradation
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-140/Product: lysozyme A #status predicted <MAT>

Query Match 86.8%; Score 33; DB 2; Length 140;
 Best Local Similarity 50.0%; Pred. No. 62;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 123 WSTWHY 128

RESULT 34

S41574
 Lysozyme (EC 3.2.1.17) B precursor - fruit fly (Drosophila melanogaster)
 C;Species: Drosophila melanogaster
 C;Date: 25-Dec-1994 #sequence_revision 23-Feb-1996 #text_change 22-Jun-1999
 C;Accession: S41574; S32643; S32620
 R;Daffre, S.; Kilsten, P.; Samakovlis, C.; Hultmark, D.
 Mol. Gen. Genet. 242, 152-162, 1994
 A;Title: The lysozyme locus in Drosophila melanogaster: an expanded gene family adapte
 A;Reference number: S41573; MUID:94211204; PMID:8159165
 A;Accession: S41574
 A;Molecule type: mRNA
 A;Residues: 1-140 <DAF>
 A;Cross-references: EMBL:Z22225; NID:g289001; PIDN:CAA80227.1; PID:g289002
 C;Genetics: LSYB

R;Daffre, S.; Kysten, P.; Samakovlis, C.; Hultmark, D.
 Submitted to the EMBL Data Library, April 1993

A;Description: The lysozyme locus in *Drosophila melanogaster*; an expanded gene family ad

A;Reference number: S32620

A;Accession: S32643

A;Molecule type: DNA

A;Residues: 1-140 <DA2>

A;Cross-references: EMBL:Z22226

C;Genetics: LSYC

C;Genetics: <LYSE>

A;Gene: LysB

A;Cross-references: FlyBase:FBgn0004425

A;Map position: 3

C;Genetics: <LYSC>

A;Gene: FlyBase:LysC

A;Cross-references: FlyBase:FBgn0004426

A;Map position: 3

C;Superfamily: lysozyme c

C;Keywords: bacteriolytic enzyme; glycosidase; hydrolase; polysaccharide degradation

F1-18/Domain: signal sequence #status predicted <SIG>

F19-140/Product: lysozyme B #status predicted <MAT>

Query Match 86.8%; Score 33; DB 2; Length 140;

Best Local Similarity 50.0%; Pred. No. 62;

Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 123 WSTWHY 128

RESULT 35

S41577

Lysozyme (EC 3.2.1.17) E precursor - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 25-Dec-1994 #sequence_revision 23-Feb-1996 #text_change 22-Jun-1999

C;Accession: S41577; S32634

R;Daffre, S.; Kysten, P.; Samakovlis, C.; Hultmark, D.

Mol. Gen. Genet. 242, 152-162, 1994

A;Title: The lysozyme locus in *Drosophila melanogaster*: an expanded gene family adapted

A;Reference number: S41573; MUID:94211204; PMID:8159165

A;Accession: S41577

A;Molecule type: mRNA

A;Residues: 1-140 <DAF>

A;Cross-references: EMBL:Z22227; NID:G289003; PIDN:CAA80229.1; PID:G289004

C;Genetics: LysE

A;Gene: LysE

A;Cross-references: FlyBase:FBgn0004428

A;Map position: 3

C;Superfamily: lysozyme c

C;Keywords: bacteriolytic enzyme; glycosidase; hydrolase; polysaccharide degradation

F1-18/Domain: signal sequence #status predicted <SIG>

F19-140/Product: lysozyme E #status predicted <MAT>

Query Match 86.8%; Score 33; DB 2; Length 140;

Best Local Similarity 50.0%; Pred. No. 62;

Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 123 WSTWHY 128

RESULT 36

B71353

Probable D,D-carboxypeptidase - syphilis spirochete

C;Species: *Treponema pallidum* subsp. *pallidum* (syphilis spirochete)

C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999

C;Accession: B71353

R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

rson, J.; Khaliq, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A;Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.

A;Reference number: A71250; MUID:98332770; PMID:9665876

A;Accession: B71353

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-260 <COL>

A;Cross-references: GB:AE001203; GB:AE000520; NID:G3322476; PIDN:AA65208.1; PID:G332

A;Experimental source: strain Nichols

C;Genetics:

A;Gene: TP0221

Query Match 86.8%; Score 33; DB 2; Length 260;

Best Local Similarity 66.7%; Pred. No. 1.1e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 220 WEPWHF 225

RESULT 37

A83299

Hypothetical protein PA2778 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C;Species: *Pseudomonas aeruginosa*

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C;Accession: A83299

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: A83299

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-293 <STO>

A;Cross-references: GB:AE004705; GB:AE004091; NID:G9948851; PIDN:AAG06166.1; GSPDB:GN

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA2778

Query Match 86.8%; Score 33; DB 2; Length 293;

Best Local Similarity 66.7%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6

Db 101 WPRWHF 106

RESULT 38

B87540

Hypothetical protein CC2347 [imported] - *Caulobacter crescentus*

C;Species: *Caulobacter crescentus*

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C;Accession: B87540

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: B87540

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-488 <STO>

A;Cross-references: GB:AE005673; NID:G13423874; PIDN:AAK24318.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC2347

C;Superfamily: Escherichia coli probable transport protein b0511

Query Match 86.8%; Score 33; DB 2; Length 488;

Best Local Similarity 66.7%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
|
23 WXXWHF 28

Db

RESULT 39
AH2079
alpha-amylase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AH2079
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2079
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-492 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA073889.1; PID:G17131281; GSPDB:GN001179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2190
C:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology

Query Match 86.8%; Score 33; DB 2; Length 492;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
|
157 WHWWHF 162

Db

RESULT 40
T02790
hypothetical protein L549.2 [imported] - Leishmania major (strain Friedlin)
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C:Accession: B81455; T02790
R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A:Reference number: B81455; MUID:99178987; PMID:10077609
A:Accession: B81455
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-549 <PYL>
A:Cross-references: GB:AB001274; NID:G3264850; PIDN:AAC24614.1; PID:G2978451; GSPDB:GN00
A:Experimental source: strain MHOM/IL/81/Friedlin
C:Genetics:
A:Gene: L549.2
A:Map position: 1
C:Superfamily: Leishmania major probable membrane protein L549.2
C:Keywords: transmembrane protein

Query Match 86.8%; Score 33; DB 2; Length 549;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
|
353 WTAWHY 358

Db

RESULT 41
T28631
Y4cD protein - Rhizobium sp. plasmid pNGR234a
C:Species: Rhizobium sp.

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T28631
R:Freiberg, C.; Fellay, R.; Bairoch, A.; Broughton, W.J.; Rosenthal, A.; Perret, X.
Nature 387, 394-401, 1997
A:Title: Molecular basis of symbiosis between Rhizobium and legumes.
A:Reference number: Z14734; MUID:97305956; PMID:9163424
A:Accession: T28631
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-640 <FRE>
A:Cross-references: EMBL:AE000067; NID:G2182324; PID:G2182328; PIDN:AAB91634.1
A:Experimental source: strain NGR234
C:Genetics:
A:Gene: Y4cD
A:Genome: plasmid pNGR234a
C:Superfamily: Rhizobium plasmid pNGR234a Y4cD protein

Query Match 86.8%; Score 33; DB 2; Length 640;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
|
521 WPFWHF 526

Db

RESULT 42
T29448
hypothetical protein F08F3.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 12-Jun-2003
C:Accession: T29448
R:Blanchard, M.; Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid F08F3.
A:Reference number: Z20620
A:Accession: T29448
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-718 <BLA>
A:Cross-references: EMBL:U64847; PIDN:AAB04876.1; GSPDB:GN00023; CESP:F08F3.2
A:Experimental source: strain Bristol N2; clone F08F3
C:Genetics:
A:Gene: CESP:F08F3.2
A:Map position: 5
A:Incons: 42/3; 45/2; 156/3; 279/1; 310/3; 346/3; 406/1; 516/1; 552/3; 600/3; 667/1
C:Superfamily: glycerol-3-phosphate O-acyltransferase

Query Match 86.8%; Score 33; DB 2; Length 718;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
|
178 WCNWHF 183

Db

RESULT 43
AG1978
hypothetical protein alr1378 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AG1978
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguc
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG1978
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-779 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA073335.1; PID:g17130725; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1378

Query Match 86.8%; Score 33; DB 2; Length 779;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
| | | |
Db 142 WCLWHF 147

RESULT 44
S41579
Lysozyme (EC 3.2.1.17) S precursor - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C>Date: 25-Dec-1994 #sequence_revision 23-Feb-1996 #text_change 22-Jun-1999
C:Accession: S41579; S32635
R:Daffre, S.; Kylisten, P.; Samakovlis, C.; Hultmark, D.
Mol. Gen. Genet. 242, 152-162, 1994
A:Title: The lysozyme locus in *Drosophila melanogaster*: an expanded gene family adapted
A:Reference number: S41573; MUID:94211204; PMID:8159165
A:Accession: S41579
A:Molecule type: mRNA
A:Residues: 1-139 <DAF>
A:Cross-references: EMBL:Z22228; NID:G289005; PIDN:CAA80230.1; PID:G289006
C:Genetics:
A:Gene: lyss
A:Cross-references: FlyBase:FBgn004430
A:Map position: 3
C:Superfamily: lysozyme c
C:Keywords: bacteriolytic enzyme; glycosidase; hydrolase; polysaccharide degradation
F:1-18/Domain: signal sequence #status predicted <STG>
F:19-139/Product: lysozyme S #status predicted <MAT>

Query Match 84.2%; Score 32; DB 2; Length 139;
Best Local Similarity 50.0%; Pred. No. 90;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
| | | |
Db 122 WAWWHY 127

RESULT 45
T01761
Hypothetical protein A_IG002P16.5 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 22-Oct-1999
C:Accession: T01761
R:Miller, N.; Beck, C.; Kramer, J.
Submitted to the EMBL Data Library, June 1997
A:Description: The sequence of A. thaliana IG002P16.
A:Reference number: Z14421
A:Accession: T01761
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-244 <MIL>
A:Cross-references: EMBL:AF007270; NID:G2191157; PID:G2191160; GSPDB:GN00063; ATSP:A_IG0
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:A_IG002P16.5
A:Map position: 5
A:Introns: 71/3; 149/1; 190/2

Query Match 84.2%; Score 32; DB 2; Length 244;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
| | | |

Db 181 WYSWHY 186

RESULT 46

D70753
Probable oxidoreductase - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70753
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70753
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-372 <COL>
A:Cross-references: GB:Z77137; GB:AL123456; NID:G3261593; PIDN:CAB00893.1; PID:e25494
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1260

Query Match 84.2%; Score 32; DB 2; Length 372;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
| | | |
Db 195 WQTHY 200

RESULT 47

B64987
bicyclomycin resistance protein - *Escherichia coli* (strain K-12)
C:Species: *Escherichia coli*
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C:Accession: B64987; JN0659
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64987
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <BLAT>
A:Cross-references: GB:AE000308; GB:U00096; NID:G1789508; PIDN:AACT5243.1; PID:g17895
A:Experimental source: strain K-12, substrain MG1655
R:Bentley, J.; Hyatt, L.S.; Ainley, K.; Parish, J.H.; Herbert, R.B.; White, G.R.
Gene 127, 117-120, 1993
A:Title: Cloning and sequence analysis of an *Escherichia coli* gene conferring bicyclor
A:Reference number: JN0659; MUID:93252267; PMID:8486276
A:Accession: JN0659
A:Molecule type: DNA
A:Residues: 20-98, 'V', 100-245, 'I', 247-396 <BEN>
A:Cross-references: EMBL:X63703; NID:G41064; PIDN:CAA45230.1; PID:g41065
C:Genetics:
A:Gene: bcr
C:Superfamily: bicyclomycin resistance protein
C:Keywords: antibiotic resistance

Query Match 84.2%; Score 32; DB 1; Length 396;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
| | | |
Db 160 WLSWHY 165

RESULT 48

B91013
 bicyclomycin resistance protein [imported] - Escherichia coli (strain O157:H7, substrain B91013)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: B91013
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome reference number: A99629; PMID:11258796
 A:Reference number: A99629; PMID:11258796
 A:Accession: B91013
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-396 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA36497.1; PID:gl3362543; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 050952
 C:Genetics:
 A:Gene: ECS3074
 C:Superfamily: bicyclomycin resistance protein

Query Match 84.2%; Score 32; DB 2; Length 396;
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 160 WLSWHY 165

RESULT 49

D85857
 bicyclomycin resistance protein [imported] - Escherichia coli (strain O157:H7, substrain B91013)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
 C:Accession: D85857
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; PMID:21074935; PMID:11206551
 A:Accession: D85857
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-396 <STO>
 A:Cross-references: GB:AB005174; NID:gl2516514; PIDN:AAG57320.1; GSPDB:GN00145; UWGP:Z34
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: bcr
 C:Superfamily: bicyclomycin resistance protein

Query Match 84.2%; Score 32; DB 2; Length 396;
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 160 WLSWHY 165

RESULT 50

AF0785
 bicyclomycin resistance protein [imported] - Salmonella enterica subsp. enterica serovar
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A>Note: this species has also been called Salmonella typhi
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AF0785
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar

A:Reference number: AB0502; PMID:21534947; PMID:11677608

A:Accession: AF0785
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-396 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD02604.1; PID:gl6503460; GSPDB:GN00176
 C:Genetics:
 A:Gene: bcr
 C:Superfamily: bicyclomycin resistance protein

Query Match 84.2%; Score 32; DB 2; Length 396;
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | | | |
 Db 160 WLSWHY 165

Search completed: June 10, 2004, 10:51:26
 Job time : 14.8667 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:36:54 ; Search time 8 seconds
(without alignments)
39.053 Million cell updates/sec

Title: US-09-912-414-11

Perfect score: 38
Sequence: 1 WXXWHF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	92.1	358	1 MTC1 CAUCR	Q45971 caulobacter
2	35	92.1	376	1 MTS1 RHINE	O30569 rhizobium m
3	35	92.1	377	1 MTB1 BRUAB	O30570 brucella ab
4	34	89.5	424	1 PD6C SOYBN	P48628 glycine max
5	34	89.5	439	1 Y412 ARATH	O04658 arabidopsis
6	34	89.5	443	1 PD6C BRANA	P48627 brassica na
7	34	89.5	448	1 PD6C ARATH	P46312 arabidopsis
8	34	89.5	479	1 CATA_PSEPU	Q59714 pseudomonas
9	34	89.5	504	1 SIK1 YEAST	Q12460 saccharomyc
10	34	89.5	511	1 NOP5 YEAST	Q12499 saccharomyc
11	34	89.5	529	1 NOP5 HUMAN	O92493 homo sapien
12	34	89.5	534	1 NOP5 RAT	O92486 rattus norv
13	33	86.8	140	1 LYSA_DROME	P37157 drosophila
14	33	86.8	140	1 LYSE_DROME	Q08694 drosophila
15	33	86.8	140	1 LYSE_DROME	P37159 drosophila
16	33	86.8	142	1 LYSE_DROME	P37161 drosophila
17	33	86.8	640	1 Y4CD RHISN	P55386 rhizobium s
18	33	86.8	718	1 PLSB CAHEL	Q22949 caenorhabdi
19	32	84.2	140	1 LYSS_DROME	P37160 drosophila
20	32	84.2	372	1 YC60 MYCTU	Q11058 mycobacteri
21	32	84.2	396	1 BCR_ECOLI	P28246 escherichia
22	32	84.2	483	1 SHT3 RAT	P35563 rattus norv
23	32	84.2	487	1 SHT3 MOUSE	P35563 rattus norv
24	32	84.2	490	1 SHT3 CAVPO	O70212 cavia porce
25	31	81.6	475	1 YIEO ECOLI	P37147 escherichia
26	31	81.6	482	1 CATA_ONCVE	Q27710 onchocerca
27	30	78.9	381	1 T10B MOUSE	Q92484 mus musculu
28	30	78.9	432	1 FIBG PETMA	P04115 petromyzon
29	30	78.9	554	1 PYRG CORGL	Q8917 corynebacte
30	30	78.9	883	1 RPOL BPT7	P00573 bacterioph
31	30	78.9	1028	1 FDGX HAFIN	P46448 haemophilu
32	30	78.9	3341	1 POLG_MCPA	P33515 m genome po
33	29	76.3	85	1 VG73_BPML5	Q05288 mycobacteri

34	29	76.3	112	1 VPX_HV2D1	P17760 human immun
35	29	76.3	164	1 YP98 MYCTU	Q50623 mycobacteri
36	29	76.3	170	1 LSPA SERMA	O52213 serratia ma
37	29	76.3	172	1 YDEI ECOLI	P31131 escherichia
38	29	76.3	226	1 YC07 METUA	Q58604 methanococc
39	29	76.3	228	1 LIEB_THETN	Q87960 thermoanaer
40	29	76.3	254	1 HIS4 SYNPF	Q89100 synchococc
41	29	76.3	302	1 SIC2 HUMAN	O75897 homo sapien
42	29	76.3	312	1 FGL1 HUMAN	Q08830 homo sapien
43	29	76.3	316	1 TYSY_LACCA	P00469 lactobacill
44	29	76.3	391	1 LYC1_YARLI	P41929 yarrowia li
45	29	76.3	485	1 ALGI_PSEPK	Q88nd2 pseudomonas
46	29	76.3	495	1 ALGI_PSEEL	P53789 pseudomonas
47	29	76.3	499	1 ALGI_AZOVI	O52196 azotobacter
48	29	76.3	518	1 ALGI_PSEEM	Q88766 pseudomonas
49	29	76.3	520	1 ALGI_PSEAB	Q51392 pseudomonas
50	29	76.3	550	1 PM21_LYCHS	P09607 lycopersico
51	29	76.3	560	1 GUP1 YEAST	Q825r0 salmonella
52	29	76.3	564	1 YEDQ_ECO57	Q825r0 salmonella
53	29	76.3	564	1 YEDQ_ECOLI	Q825r0 salmonella
54	29	76.3	564	1 YEDQ_SALTI	Q825r0 salmonella
55	29	76.3	570	1 YEN1_SCHPO	Q825r0 salmonella
56	29	76.3	572	1 YEN1_SCHPO	Q825r0 salmonella
57	29	76.3	588	1 YAT1_SCHPO	Q825r0 salmonella
58	29	76.3	609	1 GUP2 YEAST	Q825r0 salmonella
59	29	76.3	627	1 FLGK BORBU	Q825r0 salmonella
60	29	76.3	815	1 YBQ2 ECOLI	Q825r0 salmonella
61	29	76.3	915	1 YBQ2 ECOLI	Q825r0 salmonella
62	29	76.3	967	1 YBQ2 ECOLI	Q825r0 salmonella
63	29	76.3	974	1 YBQ2 ECOLI	Q825r0 salmonella
64	29	76.3	986	1 YBQ2 ECOLI	Q825r0 salmonella
65	29	76.3	1015	1 YBQ2 ECOLI	Q825r0 salmonella
66	29	76.3	1016	1 YBQ2 ECOLI	Q825r0 salmonella
67	29	76.3	1021	1 YBQ2 ECOLI	Q825r0 salmonella
68	29	76.3	1285	1 TOXA_PASNU	Q825r0 salmonella
69	29	76.3	1293	1 MLE_DROME	Q825r0 salmonella
70	29	76.3	1316	1 N160_HUMAN	Q825r0 salmonella
71	29	76.3	1402	1 N160_HUMAN	Q825r0 salmonella
72	29	76.3	3412	1 POLG_TBTVS	Q825r0 salmonella
73	29	76.3	3414	1 POLG_TBTVS	Q825r0 salmonella
74	29	76.3	3414	1 POLG_TBTVH	Q825r0 salmonella
75	29	76.3	3414	1 POLG_TBTVH	Q825r0 salmonella
76	29	76.3	3415	1 POLG_TBTVH	Q825r0 salmonella
77	29	76.3	3415	1 POLG_TBTVH	Q825r0 salmonella
78	29	76.3	4543	1 LRPI_HUMAN	Q825r0 salmonella
79	29	76.3	4544	1 LRPI_HUMAN	Q825r0 salmonella
80	29	76.3	79	1 YKFF ECOLI	Q825r0 salmonella
81	29	76.3	90	1 YKFF ECOLI	Q825r0 salmonella
82	29	76.3	103	1 ES6D MYCTU	Q825r0 salmonella
83	29	76.3	105	1 NIGM_HUMAN	Q825r0 salmonella
84	29	76.3	108	1 NIGM_HUMAN	Q825r0 salmonella
85	29	76.3	111	1 VPX_HV2KR	Q825r0 salmonella
86	29	76.3	111	1 YFBW_SALTY	Q825r0 salmonella
87	29	76.3	111	1 VPX_HV2CA	Q825r0 salmonella
88	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
89	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
90	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
91	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
92	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
93	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
94	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
95	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
96	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
97	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
98	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
99	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
100	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
101	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
102	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
103	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
104	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
105	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella
106	29	76.3	112	1 VPX_HV2G1	Q825r0 salmonella

107	28	73.7	260	1	ERS1_YEAST	P17261	saccharomyc	180	28	73.7	1034	1	BGAL_KLEPN	P06219	klebsiella
108	28	73.7	260	1	RPBA_MYXXA	Q50862	myxococcus	181	28	73.7	1049	1	SPS_ORISA	Q43802	oryza sativ
109	28	73.7	267	1	COML_NEIGO	Q50985	neisseria g	182	28	73.7	1063	1	DPOM_CLAPU	P22373	claviceps p
110	28	73.7	267	1	COML_NEIMA	Q9JVB7	neisseria m	183	28	73.7	1068	1	SPS_MAIZE	P31927	zea mays (m
111	28	73.7	267	1	COML_NEIMA	Q9KUB1	neisseria m	184	28	73.7	1081	1	SPS2_CRAPL	O04933	craterostig
112	28	73.7	287	1	YAHE_ECOLI	P77297	escherichia	185	28	73.7	1169	1	SUV3_DROME	P20193	drosophila
113	28	73.7	300	1	YIIP_ECOLI	P32159	escherichia	186	28	73.7	1193	1	ACE_CHICK	Q10751	gallus gall
114	28	73.7	310	1	ARAC_BRWCH	P07482	erwinia chr	187	28	73.7	1193	1	DPOL_ADE04	P87503	human adeno
115	28	73.7	316	1	CEL_ECOLI	Q47043	escherichia	188	28	73.7	1197	1	DPOM_PODAN	Q01529	podospora a
116	28	73.7	316	1	CEL_KLEAE	Q08598	klebsiella	189	28	73.7	1281	1	YLB5_CAEEL	P46580	caenorhabdi
117	28	73.7	320	1	XERC_SINP7	Q8KUV2	synecococc	190	28	73.7	1306	1	ACE_HUMAN	P12821	homo sapien
118	28	73.7	343	1	Y098_GVCL	P41729	cryptophleb	191	28	73.7	1310	1	ACE_RABIT	P12822	oryctolagus
119	28	73.7	350	1	EUTR_ECOLI	P36547	escherichia	192	28	73.7	1312	1	ACE_MOUSE	P09470	mus muscucu
120	28	73.7	350	1	EUTR_SALTY	Q9ZFU7	salmonella	193	28	73.7	1313	1	ACE_RAT	P47820	rattus norv
121	28	73.7	363	1	RL4A_SCHPO	P35679	schizosacch	194	28	73.7	1333	1	ELF1_DROME	P13002	drosophila
122	28	73.7	363	1	RL4B_SCHPO	Q9P784	schizosacch	195	28	73.7	1520	1	ACFD_ECOLI	Q46837	escherichia
123	28	73.7	379	1	O3B3_DROME	P81915	drosophila	196	28	73.7	1520	1	ACFD_VIBCH	Q9RTQ4	vibrio chol
124	28	73.7	382	1	NU2M_CHLRE	P08740	chlamydomon	197	28	73.7	1698	1	CUL7_HUMAN	Q14999	homo sapien
125	28	73.7	384	1	A2AB_ECHTE	O77723	echinops te	198	28	73.7	3660	1	DMD_CHICK	P11533	gallus gall
126	28	73.7	386	1	RL4_URECA	P49165	urechis cau	199	28	73.7	3678	1	DMD_MOUSE	P11531	mus muscucu
127	28	73.7	394	1	FTSW_HAEIN	P45064	haemophilus	200	28	73.7	3680	1	DMD_CANFA	O97592	canis famil
128	28	73.7	395	1	DLTB_BACSU	P39580	bacillus eu	201	28	73.7	3685	1	DMD_HUMAN	P11532	homo sapien
129	28	73.7	396	1	RL4A_XENLA	P08429	xenopus lae	202	27	71.1	148	1	RS12_METTA	P54062	methanococc
130	28	73.7	396	1	RL4B_XENLA	P02385	xenopus lae	203	27	71.1	171	1	Y428_MYCGB	P47667	mycoplasma
131	28	73.7	400	1	ASSY_CLOAB	Q97KE6	clostridium	204	27	71.1	172	1	YB03_MYCPN	P75566	mycoplasma
132	28	73.7	406	1	RL4B_ARATH	Q9SF40	arabidopsis	205	27	71.1	181	1	AA02_MYCTU	P95219	mycobacteri
133	28	73.7	407	1	RL4A_ARATH	P49691	arabidopsis	206	27	71.1	256	1	Y090_MYCTU	O10887	mycobacteri
134	28	73.7	408	1	RL4A_ARATH	Q9AF97	prunus arme	207	27	71.1	271	1	YD17_HAEIN	P44160	haemophilus
135	28	73.7	409	1	POBA_PSEPS	Q5185	pseudomonas	208	27	71.1	287	1	LEP4_LEGPN	O88433	legionella
136	28	73.7	419	1	RL4_MOUSE	Q948e6	mus muscucu	209	27	71.1	318	1	Y501_SINP3	Q55487	synecocyst
137	28	73.7	421	1	RL4_MOUSE	Q28346	canis famil	210	27	71.1	320	1	ASPG_SPOFR	O2467	spodoptera
138	28	73.7	421	1	PNK1_SCHPO	O13911	schizosacch	211	27	71.1	322	1	NU1M_PELSU	O79670	pelomedusa
139	28	73.7	421	1	RL4_RAT	P50878	rattus norv	212	27	71.1	380	1	CYB_RANCA	P16674	rana catesb
140	28	73.7	427	1	RL4_HUMAN	P35578	homo sapien	213	27	71.1	382	1	YCAD_ECOLI	P21603	escherichia
141	28	73.7	433	1	YBBY_ECOLI	P77328	escherichia	214	27	71.1	413	1	ACKA_SINP3	P21503	escherichia
142	28	73.7	435	1	Y990_MYCTU	P71771	mycobacteri	215	27	71.1	433	1	RSTB_ECOLI	P73152	synecocyst
143	28	73.7	440	1	SCRC_HUMAN	P47872	homo sapien	216	27	71.1	439	1	YSM3_CABEL	P18392	escherichia
144	28	73.7	445	1	SCRC_RABIT	O46502	oryctolagus	217	27	71.1	482	1	NF31_NAEFO	Q10123	caenorhabdi
145	28	73.7	447	1	VIPIR_CARAU	Q90308	carassius a	218	27	71.1	508	1	MLO3_ARATH	P42661	naegleria f
146	28	73.7	449	1	SCRC_RAT	P32811	rattus norv	219	27	71.1	510	1	CBP1_ORISA	Q94KB9	arabidopsis
147	28	73.7	451	1	PKTC_ECOLI	Q93811	rattus norv	220	27	71.1	532	1	IPA7_SHIFL	P37890	oryza sativ
148	28	73.7	452	1	AMP_HUMAN	P37189	escherichia	221	27	71.1	560	1	NMB_HUMAN	P18014	shigella fl
149	28	73.7	452	1	AMP_HUMAN	Q33685	homo sapien	222	27	71.1	574	1	IPA4_SHIFL	Q14956	homo sapien
150	28	73.7	457	1	VIPIR_HUMAN	P32241	homo sapien	223	27	71.1	637	1	SKB1_HUMAN	P18009	shigella fl
151	28	73.7	457	1	VIPIR_MELGA	Q91085	meleagris g	224	27	71.1	637	1	SKB1_MOUSE	O14744	homo sapien
152	28	73.7	458	1	VIPIR_PIG	Q28992	sus acrofa	225	27	71.1	640	1	SYTM_PODAN	Q8CIG8	mus muscucu
153	28	73.7	459	1	VIPIR_MOUSE	P97751	mus muscucu	226	27	71.1	695	1	RBT1_MOUSE	P28669	podospora a
154	28	73.7	459	1	VIPIR_RAT	P30083	rattus norv	227	27	71.1	773	1	PAC2_PSES3	Q9DAK3	mus muscucu
155	28	73.7	461	1	FUCO_HUMAN	P40666	homo sapien	228	27	71.1	803	1	OPGH_PSES3	P15558	pseudomonas
156	28	73.7	462	1	FUCO_RAT	P17164	rattus norv	229	27	71.1	821	1	TRKB_RAT	P20401	pseudomonas
157	28	73.7	465	1	FUCO_CANFA	P48300	canis famil	230	27	71.1	857	1	OPGH_PSEPK	Q63604	rattus norv
158	28	73.7	475	1	ZCH5_HUMAN	Q8N8U3	homo sapien	231	27	71.1	859	1	OPGH_PSESM	Q88D04	pseudomonas
159	28	73.7	483	1	COA2_BPPF3	P03624	bacterioph	232	27	71.1	861	1	OPGH_PSEAE	Q87UY1	pseudomonas
160	28	73.7	485	1	YUDL_ECOLI	P39276	escherichia	233	27	71.1	862	1	OPGH_RALSO	Q8HUA6	pseudomonas
161	28	73.7	489	1	SVV_TRIVA	P46216	trichomonas	234	27	71.1	918	1	PEP3_YEAST	Q8XVC2	raletonia s
162	28	73.7	495	1	ROFN_ASPNG	P46882	aspergillus	235	27	71.1	1037	1	ATCG_SCHPO	P27801	saccharomyc
163	28	73.7	606	1	PF20_CHLRE	P93107	chlamydomon	236	27	71.1	1037	1	MYL5_MOUSE	R22189	schizosacch
164	28	73.7	614	1	FEOB_SINP3	P73182	synecocyst	237	27	71.1	1037	1	YGS6_MOUSE	Q9QZ24	mus muscucu
165	28	73.7	634	1	GLGE_LACPL	Q890J1	lactobacill	238	26	68.4	101	1	YGS6_MOUSE	P75147	mycoplasma
166	28	73.7	686	1	CHEA_RHOSH	Q53135	rhodobacter	239	26	68.4	104	1	YJA7_YEAST	P47080	saccharomyc
167	28	73.7	709	1	SULL_CAREL	Q21376	caenorhabdi	240	26	68.4	129	1	YHM2_YEAST	P18857	saccharomyc
168	28	73.7	754	1	YCAL_ECOLI	P37443	escherichia	241	26	68.4	138	1	YA91_MYCPN	P75602	mycoplasma
169	28	73.7	788	1	BSCB_XANAC	P58933	xanthomonas	242	26	68.4	139	1	YE13_MYCPN	Q9EXD6	mycoplasma
170	28	73.7	821	1	TRKB_MOUSE	P15209	mus muscucu	243	26	68.4	146	1	YB63_MYCPN	P53311	saccharomyc
171	28	73.7	822	1	TRKB_HUMAN	Q16620	homo sapien	244	26	68.4	157	1	YB63_MYCPN	P75320	mycoplasma
172	28	73.7	826	1	R1R1_EBV	P03190	epstein-bar	245	26	68.4	182	1	VG37_BPMU	Q9TIV8	bacterioph
173	28	73.7	969	1	DPOM_NEUIN	P33358	neurospora	246	26	68.4	213	1	IF4E_XENLA	P48597	xenopus lae
174	28	73.7	973	1	TRP5_HUMAN	Q9UL62	homo sapien	247	26	68.4	215	1	IF4E_APLCA	O77210	aplysia cal
175	28	73.7	974	1	TRP4_MOUSE	Q9QU65	mus muscucu	248	26	68.4	217	1	IF4E_BOVIN	Q9N0T5	bos taurus
176	28	73.7	975	1	TRP5_RABIT	O62852	oryctolagus	249	26	68.4	217	1	IF4E_MOUSE	P06730	homo sapien
177	28	73.7	975	1	TRP5_MOUSE	Q9QX29	mus muscucu	250	26	68.4	217	1	IF4E_MOUSE	P20415	mus muscucu
178	28	73.7	977	1	TRP4_HUMAN	Q9UBN4	homo sapien	251	26	68.4	251	1	IF4E_RABIT	P29338	oryctolagus
179	28	73.7	981	1	TRP4_RAT	Q35119	rattus norv	252	26	68.4	254	1	YABI_ECOLI	P30149	escherichia
						P75100	bos taurus				257	1	FOL1_HUMAN	P15328	homo sapien

253	26	68.4	289	1	LEP4_PSEST	Q9zel6 pseudomonas	325	25	65.8	122	1	BMG_ACIBB	Q9prf8 acipenser b
254	26	68.4	290	1	LEP4_PSEAE	P22610 pseudomonas	327	25	65.8	123	1	CD59_PIG	O62680 sus scrofa
255	26	68.4	307	1	TR41_HUMAN	P59536 homo sapien	328	25	65.8	145	1	RL32_AERPE	Q9yf92 aetopyrum p
256	26	68.4	307	1	TR59_HUMAN	P59536 homo sapien	329	25	65.8	146	1	KLA2_ECOLI	P52603 escherichia
257	26	68.4	308	1	T2RC_MOUSE	P59532 mus musculus	330	25	65.8	146	1	PA2A_NAJSP	Q92084 najia sputat
258	26	68.4	308	1	T2RC_MOUSE	Q9jke7 rattus norv	331	25	65.8	146	1	PA2B_NAJSP	Q92085 najia sputat
259	26	68.4	313	1	MTH1_HAEP4	P29538 haemophilus	332	25	65.8	147	1	COX3_SPOPR	Q38226 spodoptera
260	26	68.4	333	1	PFTA_PEA	O24304 p protein f	333	25	65.8	147	1	CST9_HUMAN	Q9h491 homo sapien
261	26	68.4	340	1	PFTA_ARATH	Q91x33 a protein f	334	25	65.8	147	1	Y565_METUA	Q57985 methanococ
262	26	68.4	340	1	PFTA_BOVIN	P29702 b protein f	335	25	65.8	153	1	HOPD_ECOLI	O842r1 wigglewort
263	26	68.4	348	1	GAL7_ECOLI	P09148 escherichia	336	25	65.8	155	1	LSPA_WIGBR	O86932 escherichia
264	26	68.4	348	1	GAL7_ECOLI	P22714 salmonella	337	25	65.8	156	1	LSPA_BUCAP	Q8k9z3 buchnera ap
265	26	68.4	350	1	GAL7_DROME	Q9vma2 drosophila	338	25	65.8	160	1	LSPA_BUCBP	P57248 buchnera ap
266	26	68.4	370	1	TAM2_HUMAN	Q15035 homo sapien	339	25	65.8	163	1	LSPA_ECOLI	P57248 buchnera ap
267	26	68.4	372	1	Y906_MYCTU	Q10562 mycobacteri	340	25	65.8	164	1	LSPA_ECOLI	P57248 buchnera ap
268	26	68.4	377	1	PFTA_MOUSE	O61239 m protein f	341	25	65.8	164	1	LSPA_ECOLI	Q8k9z3 buchnera ap
269	26	68.4	377	1	PFTA_MOUSE	Q04631 r protein f	342	25	65.8	164	1	LSPA_ECOLI	Q8k9z3 buchnera ap
270	26	68.4	377	1	PFTA_MOUSE	Q49354 h protein f	343	25	65.8	164	1	LSPA_ECOLI	Q8k9z3 buchnera ap
271	26	68.4	379	1	PFTA_MOUSE	P40908 cryptococcu	344	25	65.8	165	1	LSPA_ENTAE	P10804 escherichia
272	26	68.4	381	1	GAL7_CHLYE	P20113 chlamydomon	345	25	65.8	166	1	LSPA_SALTI	P13514 enterobacte
273	26	68.4	443	1	NU4M_CHLYE	O06127 mycobacteri	346	25	65.8	166	1	LSPA_SALTI	Q829n1 salmonella
274	26	68.4	516	1	TRPE_MYCTU	Q9x7c5 mycobacteri	347	25	65.8	166	1	LSPA_SALTI	Q829n1 salmonella
275	26	68.4	529	1	TRPE_MYCTU	P15878 mycobacteri	348	25	65.8	169	1	LSPA_SALTI	Q829n1 salmonella
276	26	68.4	551	1	QCRB_MYCLE	Q10279 schizosacch	349	25	65.8	176	1	YMI6_ARCFU	Q829n1 salmonella
277	26	68.4	581	1	FUR4_SCHPO	Q10279 schizosacch	350	25	65.8	176	1	YMI6_ARCFU	Q829n1 salmonella
278	26	68.4	581	1	ZDS_TARER	Q9fv46 tagetes ere	351	25	65.8	178	1	COXZ_RICPR	P26701 bacterioph
279	26	68.4	641	1	MLA2_HUMAN	O60476 homo sapien	352	25	65.8	179	1	RL5_BUCAL	Q9adms rickettsia
280	26	68.4	641	1	MLA2_MOUSE	P39098 mus musculu	353	25	65.8	180	1	YCBW_ECOLI	P57579 buchnera ap
281	26	68.4	647	1	AGSA_CAUCR	Q9a2i0 caulobacter	354	25	65.8	182	1	YDJA_ECOLI	P75862 escherichia
282	26	68.4	648	1	AMOI_ARTS1	Q07121 arthrobacte	355	25	65.8	183	1	YDJA_ECOLI	Q57457 haemophilus
283	26	68.4	657	1	CSF1_CORGL	Q07121 arthrobacte	356	25	65.8	185	1	YB06_METUA	P24250 escherichia
284	26	68.4	666	1	AMO_LENCU	Q01377 corynebacte	357	25	65.8	185	1	YB06_METUA	Q58506 methanococ
285	26	68.4	671	1	AMO_LENCU	P49252 lens culina	358	25	65.8	187	1	YB06_METUA	Q33398 escherichia
286	26	68.4	671	1	AMO_LENCU	P49252 lens culina	359	25	65.8	187	1	YB06_METUA	Q33398 escherichia
287	26	68.4	671	1	AMO_LENCU	Q12556 aspergillus	360	25	65.8	188	1	YB06_METUA	Q33398 escherichia
288	26	68.4	702	1	FOX4_SALTY	Q43077 pium sativ	361	25	65.8	188	1	YB06_METUA	Q33398 escherichia
289	26	68.4	750	1	CTPB_MYCLE	Q56145 salmonella	362	25	65.8	188	1	YB06_METUA	Q33398 escherichia
290	26	68.4	752	1	CTPB_MYCHO	P46840 mycobacteri	363	25	65.8	188	1	YB06_METUA	Q33398 escherichia
291	26	68.4	752	1	CTPB_MYCHO	P59947 mycobacteri	364	25	65.8	188	1	YB06_METUA	Q33398 escherichia
292	26	68.4	761	1	CTPA_MYCTU	Q10877 mycobacteri	365	25	65.8	189	1	YB06_METUA	Q33398 escherichia
293	26	68.4	766	1	STB6_YEAST	Q10877 mycobacteri	366	25	65.8	189	1	YB06_METUA	Q33398 escherichia
294	26	68.4	780	1	CTPA_MYCLE	P36085 saccharomyc	367	25	65.8	191	1	YB06_METUA	Q33398 escherichia
295	26	68.4	794	1	VTB_BPT7	P46839 mycobacteri	368	25	65.8	191	1	YB06_METUA	Q33398 escherichia
296	26	68.4	906	1	RIR1_HCMVA	P03747 bacterioph	369	25	65.8	191	1	YB06_METUA	Q33398 escherichia
297	26	68.4	1041	1	CHS1_CRYNE	P16782 human cytom	370	25	65.8	191	1	YB06_METUA	Q33398 escherichia
298	26	68.4	1086	1	SYI_RICPR	O13356 cryptococcu	371	25	65.8	191	1	YB06_METUA	Q33398 escherichia
299	26	68.4	1225	1	CTD2_HUMAN	Q9zcu4 rickettsia	372	25	65.8	191	1	YB06_METUA	Q33398 escherichia
300	26	68.4	1247	1	CTD2_MOUSE	Q9zcu4 rickettsia	373	25	65.8	191	1	YB06_METUA	Q33398 escherichia
301	25.5	67.1	3530	1	MY15_HUMAN	O35927 mus musculu	374	25	65.8	191	1	YB06_METUA	Q33398 escherichia
302	25.5	67.1	554	1	HUTU_CAUCR	Q9ukn7 homo sapien	375	25	65.8	191	1	YB06_METUA	Q33398 escherichia
303	25.5	67.1	555	1	HUTU_XANCP	Q9a9m1 caulobacter	376	25	65.8	191	1	YB06_METUA	Q33398 escherichia
304	25.5	67.1	557	1	HUTU_RHIME	P58988 xanthomonas	377	25	65.8	191	1	YB06_METUA	Q33398 escherichia
305	25.5	67.1	559	1	HUTU_PSEAE	Q92v80 rhizobium m	378	25	65.8	191	1	YB06_METUA	Q33398 escherichia
306	25.5	67.1	561	1	HUTU_SALTI	Q9hu83 pseudomonas	379	25	65.8	191	1	YB06_METUA	Q33398 escherichia
307	25.5	67.1	564	1	HUTU_TRIRP	O82897 salmonella	380	25	65.8	191	1	YB06_METUA	Q33398 escherichia
308	25.5	67.1	565	1	HUTU_PSESM	P53385 trifolium r	381	25	65.8	191	1	YB06_METUA	Q33398 escherichia
309	25	65.8	65	1	HA2_ECTHL	Q87um6 pseudomonas	382	25	65.8	191	1	YB06_METUA	Q33398 escherichia
310	25	65.8	82	1	TXA1_ACTEQ	Q87um6 pseudomonas	383	25	65.8	191	1	YB06_METUA	Q33398 escherichia
311	25	65.8	95	1	ESXA_MYCLE	P80103 ectothiorho	384	25	65.8	191	1	YB06_METUA	Q33398 escherichia
312	25	65.8	97	1	ATPK_YEAST	Q9njq2 actinia equ	385	25	65.8	191	1	YB06_METUA	Q33398 escherichia
313	25	65.8	101	1	ATPK_YEAST	Q50206 mycobacteri	386	25	65.8	191	1	YB06_METUA	Q33398 escherichia
314	25	65.8	104	1	Y49L_SYNY3	Q6405 cyanidium c	387	25	65.8	191	1	YB06_METUA	Q33398 escherichia
315	25	65.8	112	1	TXA1_AGEAP	Q55720 synechocyst	388	25	65.8	191	1	YB06_METUA	Q33398 escherichia
316	25	65.8	118	1	PA21_NAJME	P15969 agelenopsis	389	25	65.8	191	1	YB06_METUA	Q33398 escherichia
317	25	65.8	118	1	PA21_NAJME	P00599 najia melano	390	25	65.8	191	1	YB06_METUA	Q33398 escherichia
318	25	65.8	118	1	PA22_NAJMO	P00602 najia mossam	391	25	65.8	191	1	YB06_METUA	Q33398 escherichia
319	25	65.8	118	1	PA23_NAJMO	P00603 najia mossam	392	25	65.8	191	1	YB06_METUA	Q33398 escherichia
320	25	65.8	118	1	PA23_NAJMO	P00604 najia mossam	393	25	65.8	191	1	YB06_METUA	Q33398 escherichia
321	25	65.8	118	1	PA23_NAJMO	P00605 najia mossam	394	25	65.8	191	1	YB06_METUA	Q33398 escherichia
322	25	65.8	119	1	PA21_HEMHA	P14556 najia pallid	395	25	65.8	191	1	YB06_METUA	Q33398 escherichia
323	25	65.8	119	1	PA22_ASFSC	P00595 hemachatus	396	25	65.8	191	1	YB06_METUA	Q33398 escherichia
324	25	65.8	119	1	PA22_NAJME	P07037 aspidelaps	397	25	65.8	191	1	YB06_METUA	Q33398 escherichia
325	25	65.8	119	1	PA23_NAJME	P00600 najia melano	398	25	65.8	191	1	YB06_METUA	Q33398 escherichia
						P00601 najia melano							

Q57768 methanococ

399	25	65.8	219	1	TER3_ECOLI	P03039 escherichia	472	25	65.8	259	1	FRI4_ARATH	Q9lyn2 arabidopsis
400	25	65.8	220	1	ENVR_ECOLI	P31676 escherichia	473	25	65.8	259	1	RNP6_MOUSE	Q9c901 mus musculus
401	25	65.8	221	1	DSBI_ECOLI	Q8fid3 escherichia	474	25	65.8	260	1	COX2_BTVU	P98012 beta vulgar
402	25	65.8	222	1	DSBI_SALTY	Q8xek0 salmorella	475	25	65.8	260	1	COX2_MAIZE	P00412 zea mays (m
403	25	65.8	223	1	DODA_AAMU	P87064 amanita mus	476	25	65.8	260	1	COX2_ORYSA	P04773 oryza sativ
404	25	65.8	224	1	Y274_AQUAE	O66631 aquifex aeo	477	25	65.8	260	1	COX2_SOYBA	P05491 glycine max
405	25	65.8	225	1	ARTM_ARTSA	P17720 artemia sal	478	25	65.8	260	1	COX2_WHEAT	P00413 triticum ae
406	25	65.8	226	1	Y531_BUCAP	Q8k929 buchnera ap	479	25	65.8	260	1	COX3_ASTPE	P05505 rattus norv
407	25	65.8	227	1	NQO2_HUMAN	P16083 homo sapien	480	25	65.8	260	1	COX3_PISOC	P25003 pisaster oc
408	25	65.8	228	1	X394_PSAE	P24562 pseudomonas	481	25	65.8	260	1	COX3_RAT	Q33824 asterina pe
409	25	65.8	229	1	YLME_BACSU	O31727 bacillus su	482	25	65.8	260	1	COX3_STRPU	P05505 rattus norv
410	25	65.8	230	1	COX3_SYNY3	Q06475 synechocyst	483	25	65.8	260	1	COX3_XENLA	P15546 strongyloce
411	25	65.8	231	1	Y112_PASMU	Q8cpd5 pasteurella	484	25	65.8	260	1	MTM2_MORBO	P23192 moraxella b
412	25	65.8	232	1	YP11_VIBAL	P52055 vibrio algi	485	25	65.8	261	1	COX2_DAUCA	P27168 daucus caro
413	25	65.8	233	1	Y599_ARCFU	O28773 archaeroglob	486	25	65.8	261	1	COX3_AEPME	O47691 aepyceros m
414	25	65.8	234	1	YGG8_ECOLI	P52054 escherichia	487	25	65.8	261	1	COX3_ANTCE	O47702 antilope ce
415	25	65.8	235	1	Y461_VIBCH	Q9kuq4 vibrio chol	488	25	65.8	261	1	COX3_ANTMR	O47701 antidorcas
416	25	65.8	236	1	Y090_HABEH	P44506 haemophilus	489	25	65.8	261	1	COX3_BALMU	P41395 balenopter
417	25	65.8	237	1	YNP5_CABEL	P34558 caenorhabdi	490	25	65.8	261	1	COX3_BALPH	P24389 balenopter
418	25	65.8	238	1	Y381_TREPA	Q83396 treponema p	491	25	65.8	261	1	COX3_BOVIN	P00415 bos taurus
419	25	65.8	239	1	PSD_ZYMO	Q8X583 zymomonas m	492	25	65.8	261	1	COX3_BRARE	Q9miy4 brachydanio
420	25	65.8	240	1	VDGE_ECOLI	P52109 escherichia	493	25	65.8	261	1	COX3_CANFA	Q92261 canis famil
421	25	65.8	241	1	WFD8_HUMAN	Q8iuu0 homo sapien	494	25	65.8	261	1	COX3_CARAU	Q96133 carassius a
422	25	65.8	242	1	LECA_GRTSI	P24146 griffonia s	495	25	65.8	261	1	COX3_CRPNA	O47893 cephalophus
423	25	65.8	243	1	TRUC_HABDU	P59840 haemophilus	496	25	65.8	261	1	COX3_CRSI	O03201 ceratotheri
424	25	65.8	244	1	FMCD_BACNO	P17418 bacteroides	497	25	65.8	261	1	COX3_CHICK	P18945 gallus gall
425	25	65.8	245	1	Y016_MARPO	P38459 marchantia	498	25	65.8	261	1	COX3_COTJA	Q8sew8 coturnix co
426	25	65.8	246	1	Y068_CABEL	P52057 caenorhabdi	499	25	65.8	261	1	COX3_CROLA	P34198 crossostoma
427	25	65.8	247	1	TATC_CAMJE	P75914 escherichia	500	25	65.8	261	1	COX3_CVPCA	P15952 cyprinus ca
428	25	65.8	248	1	YGAZ_ECOLI	P76630 escherichia							
429	25	65.8	249	1	YCDX_ECOLI	P75914 escherichia							
430	25	65.8	250	1	YGAZ_ECOLI	P76630 escherichia							
431	25	65.8	251	1	YCDX_ECOLI	P75914 escherichia							
432	25	65.8	252	1	YCDX_ECOLI	P75914 escherichia							
433	25	65.8	253	1	YCDX_ECOLI	P75914 escherichia							
434	25	65.8	254	1	YCDX_ECOLI	P75914 escherichia							
435	25	65.8	255	1	YCDX_ECOLI	P75914 escherichia							
436	25	65.8	256	1	YCDX_ECOLI	P75914 escherichia							
437	25	65.8	257	1	YCDX_ECOLI	P75914 escherichia							
438	25	65.8	258	1	YCDX_ECOLI	P75914 escherichia							
439	25	65.8	259	1	YCDX_ECOLI	P75914 escherichia							
440	25	65.8	260	1	YCDX_ECOLI	P75914 escherichia							
441	25	65.8	261	1	YCDX_ECOLI	P75914 escherichia							
442	25	65.8	262	1	YCDX_ECOLI	P75914 escherichia							
443	25	65.8	263	1	YCDX_ECOLI	P75914 escherichia							
444	25	65.8	264	1	YCDX_ECOLI	P75914 escherichia							
445	25	65.8	265	1	YCDX_ECOLI	P75914 escherichia							
446	25	65.8	266	1	YCDX_ECOLI	P75914 escherichia							
447	25	65.8	267	1	YCDX_ECOLI	P75914 escherichia							
448	25	65.8	268	1	YCDX_ECOLI	P75914 escherichia							
449	25	65.8	269	1	YCDX_ECOLI	P75914 escherichia							
450	25	65.8	270	1	YCDX_ECOLI	P75914 escherichia							
451	25	65.8	271	1	YCDX_ECOLI	P75914 escherichia							
452	25	65.8	272	1	YCDX_ECOLI	P75914 escherichia							
453	25	65.8	273	1	YCDX_ECOLI	P75914 escherichia							
454	25	65.8	274	1	YCDX_ECOLI	P75914 escherichia							
455	25	65.8	275	1	YCDX_ECOLI	P75914 escherichia							
456	25	65.8	276	1	YCDX_ECOLI	P75914 escherichia							
457	25	65.8	277	1	YCDX_ECOLI	P75914 escherichia							
458	25	65.8	278	1	YCDX_ECOLI	P75914 escherichia							
459	25	65.8	279	1	YCDX_ECOLI	P75914 escherichia							
460	25	65.8	280	1	YCDX_ECOLI	P75914 escherichia							
461	25	65.8	281	1	YCDX_ECOLI	P75914 escherichia							
462	25	65.8	282	1	YCDX_ECOLI	P75914 escherichia							
463	25	65.8	283	1	YCDX_ECOLI	P75914 escherichia							
464	25	65.8	284	1	YCDX_ECOLI	P75914 escherichia							
465	25	65.8	285	1	YCDX_ECOLI	P75914 escherichia							
466	25	65.8	286	1	YCDX_ECOLI	P75914 escherichia							
467	25	65.8	287	1	YCDX_ECOLI	P75914 escherichia							
468	25	65.8	288	1	YCDX_ECOLI	P75914 escherichia							
469	25	65.8	289	1	YCDX_ECOLI	P75914 escherichia							
470	25	65.8	290	1	YCDX_ECOLI	P75914 escherichia							
471	25	65.8	291	1	YCDX_ECOLI	P75914 escherichia							

ALIGNMENTS

RESULT 1
 MTCL CAUCR STANDARD; PRT; 358 AA.
 ID MTC1 CAUCR STANDARD; PRT; 358 AA.
 AC Q45971;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Modification methylease Ccrmi (EC 2.1.1.72) (Adenine-specific
 DE methyletransferase Ccrmi) (M.Ccrmi).
 GN CCRMI OR CCRM OR CCR378.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Caulobacter.
 ON NCBI_TaxID=155892;
 RX [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CB15N / NAI1000;
 RC MEDLINE=94118303; PubMed=8289276;
 RX Zweiger G., Marczynski G., Shapiro L.;
 RT "A Caulobacter DNA methyltransferase that functions only in the
 RT predivisional cell.";
 RL J. Mol. Biol. 235:472-485(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=ATCC 19089 / CB15;
 RP MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Knouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 CC Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -!- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
 CC GATC AND CAUSES SPECIFIC METHYLATION ON A-2 ON BOTH STRANDS.
 CC CCRM-MEDIATED METHYLATION HAS IMPORTANT CELLULAR FUNCTIONS.
 CC APPEARS TO CONTRIBUTE TO THE ACCURATE CELL-CYCLE CONTROL OF DNA

```

CC REPLICATION AND CELLULAR MORPHOLOGY.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U01032; AAA18913.1; -.
DR EMBL; AE005711; AAK22365.1; -.
DR PIR; A87296; A87296.
DR PIR; S43876; S43876.
DR HSSP; P11409; 1BOO.
DR REBASE; 2539; M.CcrMI.
DR TIGR; CC0378; -.
DR InterPro; IPR002295; D21N6_mtfase.
DR InterPro; IPR001091; Met trans CN4.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PR00506; D21N6MTFRASE.
DR PRINTS; PR00508; S21N4MTFRASE.
DR PROSITE; PS00092; N6_Mtase; 1.
DR Transfaser; Methyltransferase; DNA replication; Complete proteome.
FT CONFLICT 242 242 Y -> D (IN REF. 1).
SQ SEQUENCE 358 AA; 39665 MW; 05F43266F7D4C614 CRC64;

Query Match 92.1%; Score 35; DB 1; Length 358;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 332 WTYWHF 337

RESULT 2
MTSL_RHIME
ID MTSL_RHIME STANDARD; PRT; 376 AA.
AC 030569;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Modification methylase SmeIP (EC 2.1.1.72) (Adenine-specific
DE methyltransferase SmeIP) (M.SmeI) (M.CcrMI).
GN SMeIM OR CCRM OR R09926 OR SMC00021.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RA Wright R., Stephens C., Shapiro L.;
RT "The CcrM DNA methyltransferase is widespread in the alpha subdivision
RT of proteobacteria, and its essential functions are conserved in
RT Rhizobium meliloti and Caulobacter crescentus.";
RL J. Bacteriol. 179:5869-5877(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rampsperger U.,
RA Renard C., Thebault P., Vandendol M., Weidner S., Galibert F.;

```

```

RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
CC GANTC AND CAUSES SPECIFIC METHYLATION ON A-2 ON BOTH STRANDS.
CC CCRM-MEDIATED METHYLATION HAS IMPORTANT CELLULAR FUNCTIONS.
CC APPEARS TO CONTRIBUTE TO THE ACCURATE CELL-CYCLE CONTROL OF DNA
CC REPLICATION AND CELLULAR MORPHOLOGY.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF011894; AAB71350.1; -.
DR EMBL; AL591785; CAC45498.1; -.
DR REBASE; 3264; M.SmeIP.
DR InterPro; IPR002295; D21N6_mtfase.
DR InterPro; IPR001091; Met trans CN4.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PR00506; D21N6MTFRASE.
DR PRINTS; PR00508; S21N4MTFRASE.
DR PROSITE; PS00092; N6_Mtase; 1.
DR Transfaser; Methyltransferase; DNA replication; Complete proteome.
FT CONFLICT 135 141 NPMENFK -> QPDRELQ (IN REF. 1).
FT CONFLICT 157 157 P -> A (IN REF. 1).
SQ SEQUENCE 376 AA; 41442 MW; 790DE7FE3D22900A CRC64;

Query Match 92.1%; Score 35; DB 1; Length 376;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 347 WTFWHF 352

RESULT 3
MTBL_BRUAB
ID MTBL_BRUAB STANDARD; PRT; 377 AA.
AC 030570;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Modification methylase Babi (EC 2.1.1.72) (Adenine-specific
DE methyltransferase Babi) (M.Babi) (M.CcrMI).
GN BABIM OR CCRM.
OS Brucella abortus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=2335;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2308;
RA Wright R., Stephens C., Shapiro L.;
RT "The CcrM DNA methyltransferase is widespread in the alpha subdivision
RT of proteobacteria, and its essential functions are conserved in
RT Rhizobium meliloti and Caulobacter crescentus.";
RL J. Bacteriol. 179:5869-5877(1997).
CC -!- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
CC GANTC AND CAUSES SPECIFIC METHYLATION ON A-2 ON BOTH STRANDS.
CC CCRM-MEDIATED METHYLATION HAS IMPORTANT CELLULAR FUNCTIONS.
CC APPEARS TO CONTRIBUTE TO THE ACCURATE CELL-CYCLE CONTROL OF DNA
CC REPLICATION AND CELLULAR MORPHOLOGY.

```

CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-adenosyl-L-homocysteine + DNA 6-methylaminopurine.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF011895; AAB71351.1; -.

CC REBASE; 3263; M.Babi.

DR InterPro; IPR002295; D21N6 mtfase.

DR InterPro; IPR001091; Met trans CN4.

DR InterPro; IPR002941; N6/N4 MtaB.

DR InterPro; IPR002052; N6 MtaB.

DR InterPro; IPR000051; SAM bind.

DR Pfam; PF01555; N6_N4_Mtase; 1.

DR PRINTS; PR00506; D21NGMTFRASE.

DR PRINTS; PR00508; S21N4MTFRASE.

DR PROSITE; PS00092; N6_MTASE; 1.

KW Transferase; Methyltransferase; DNA replication.

SQ SEQUENCE 377 AA; 42202 MW; 657C88A25580B39D CRC64;

Query Match 92.1%; Score 35; DB 1; Length 377;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
| | | |
Db 347 WTPWHF 352

RESULT 4

FD6C SOYBN

ID FD6C SOYBN STANDARD; PRT; 424 AA.

AC P48628;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.19.-).

OS Glycine max [Soybean].

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94345008; PubMed=8066133;
RA Hitz W.D., Carlson T.J., Booth J.R. Jr., Kinney A.J., Stecca K.L.,
RA Yadav N.S.;
RT "Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA and its expression in a cyanobacterium."
RL Plant Physiol. 105:635-641(1994).

CC -1- FUNCTION: Chloroplast omega-6 fatty acid desaturase introduces the second double bond in the biosynthesis of 16:3 and 18:3 fatty acids, important constituents of plant membranes. It is thought to use ferredoxin as an electron donor and to act on fatty acids esterified to galactolipids, sulfolipids and phosphatidylglycerol.

CC -1- PATHWAY: Polyunsaturated fatty acid biosynthesis.

CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).

CC -1- DOMAIN: The histidine box domains may contain the active site and/or be involved in metal ion binding.

CC -1- SIMILARITY: Belongs to the fatty acid desaturase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch.

CC EMBL; L29215; AAA50158.1; -.

DR PIR; T07742; T07742.

DR InterPro; IPR005804; FA_desat_fam.

DR Pfam; PF00487; FA_desaturase; 1.

DR ProDom; PD001081; FA_desat_fam; 2.

KW Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
KW transit peptide.

FT TRANSIT 1 63 CHLOROPLAST (BY SIMILARITY).

FT CHAIN 64 424 OMEGA-6 FATTY ACID DESATURASE.

FT DOMAIN 165 169 HISTIDINE BOX-1.

FT DOMAIN 201 205 HISTIDINE BOX-2.

FT DOMAIN 361 365 HISTIDINE BOX-3.

SQ SEQUENCE 424 AA; 49641 MW; 4F3DF52D4B1A2009 CRC64;

Query Match 89.5%; Score 34; DB 1; Length 424;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
| | | |
Db 251 WLMWHF 256

RESULT 5

Y412 ARATH

ID Y412 ARATH STANDARD; PRT; 439 AA.

AC O04658;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Hypothetical protein At5g27120.

OS Arabidopsis thaliana (Mouse-ear cress).

GN At5g27120 OR TM021B04.12 OR T21B4.30.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016721; PubMed=11130714;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsuno M., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Speth J., Sexton M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Mulvaney E., Ozerky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,
RA Mariensten R., McCombie W.K., Wilson R.K., Murphy G., Bancroft I.,
RA Volkert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA Ransperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dikse W., Mooijman P., Klein Lankhorst R.,
RA Weizenecker T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarreal R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K., Rudd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.,
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis thaliana".
RL Nature 408:823-826(2000).

CC -1- SIMILARITY: Belongs to the NOP5/NOP56 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AF007271; AAB61073.1; --
 DR PIR; T01807; T01807.
 DR InterPro; IPR002687; Nop.
 DR Pfam; PF01798; Nop; 1.
 DR ProDom; PD004104; Nop; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 439 AA; 47932 MW; BE0E2214F9EC4FDB CRC64;

Query Match 89.5%; Score 34; DB 1; Length 439;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 | |||
 Db 184 WFGWHF 189

RESULT 6
 FD6C BRANA STANDARD; PRT; 443 AA.
 AC P48627;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.19.-).
 OS Brassica napus (Rape).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosid II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seed;
 RX MEDLINE=94345008; PubMed=8066133;
 RA Hitz W.D., Carlson T.J., Booth J.R. Jr., Kinney A.J., Stecca K.L.,
 RA Yadav N.S.;
 RT "Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA
 RT and its expression in a cyanobacterium.";
 RL Plant Physiol. 105:635-641(1994).
 CC -!- FUNCTION: Chloroplast omega-6 fatty acid desaturase introduces the
 CC second double bond in the biosynthesis of 16:3 and 18:3 fatty
 CC acids, important constituents of plant membranes. It is thought to
 CC use ferredoxin as an electron donor and to act on fatty acids
 CC esterified to galactolipids, sulfolipids and phosphatidylglycerol.
 CC -!- PATHWAY: Polyunsaturated fatty acid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
 CC -!- DOMAIN: The histidine box domains may contain the active site
 CC and/or be involved in metal ion binding.
 CC -!- SIMILARITY: Belongs to the fatty acid desaturase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; L29214; AAA50157.1; --
 DR PIR; T08136; T08136.
 DR InterPro; IPR005804; FA desat. fam.
 DR Pfam; PF00487; FA desaturase; 1.
 DR ProDom; PD001081; FA desat. fam; 2.
 KW Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
 KW Transit peptide.
 TRANSIT 1 64 CHLOROPLAST (BY SIMILARITY).

FT CHAIN 65 443 OMEGA-6 FATTY ACID DESATURASE.
 FT DOMAIN 166 170 HISTIDINE BOX-1.
 FT DOMAIN 202 206 HISTIDINE BOX-2.
 FT DOMAIN 362 366 HISTIDINE BOX-3.
 SQ SEQUENCE 443 AA; 50755 MW; 27544B33AF2781D8 CRC64;
 Query Match 89.5%; Score 34; DB 1; Length 443;
 Best Local Similarity 66.7%; Pred. No. 38;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 | |||
 Db 252 WYNWHF 257

RESULT 7
 FD6C ARATH STANDARD; PRT; 448 AA.
 AC P46312; Q9M094;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.19.-).
 GN PAD6 OR FADC OR AT4G30950 OR F6118.140.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE OF 1-418 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=95148736; PubMed=7846158;
 RA Falcone D.L., Gibson S., Lemieux B., Somerville C.R.;
 RT "Identification of a gene that complements an Arabidopsis mutant
 RT deficient in chloroplast omega 6 desaturase activity.";
 RL Plant Physiol. 106:1453-1459(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Deisenz M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hobeisel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandebussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,
 RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Mooljman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysbaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rehnemann S.,
 RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
 RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fatmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chefdro F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Baigues M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielle C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,

RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mixx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulten L., Mardis E., Dante M., Pepin K., Hillier L.W.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M.A., Martensen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana";
 RL Nature 402:769-777 (1999).
 RN [3].
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Ejima A., Goldsmith A.D., Gurjani M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 genome";
 RL Science 302:842-846 (2003).
 CC -!- FUNCTION: Chloroplast omega-6 fatty acid desaturase introduces
 CC the second double bond in the biosynthesis of 16:3 and 18:3 fatty
 CC acids, important constituents of plant membranes. It is thought
 CC to use ferredoxin as an electron donor and to act on fatty acids
 CC esterified to galactolipids, sulfolipids and phosphatidylglycerol.
 CC -!- PATHWAY: Polyunsaturated fatty acid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
 CC -!- DEVELOPMENTAL STAGE: Highest levels found in expanding leaves.
 CC -!- DOMAIN: The histidine box domains may contain the active site
 CC and/or be involved in metal ion binding.
 CC -!- SIMILARITY: Belongs to the fatty acid desaturase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U09503; AAA92800.1; -;
 DR EMBL; AL022198; CAB18198.1; -;
 DR EMBL; AL161578; CAB9813.1; -;
 DR EMBL; AY045621; AAK73979.1; -;
 DR EMBL; AY058078; AAL24186.1; -;
 DR EMBL; AY058852; AAL24240.1; -;
 DR PIR; D85362; D85362.
 DR InterPro; IPR005804; FA_desat_fam.
 DR Pfam; PF00487; FA_desaturase; 1.
 DR ProDom; PD001081; FA_desat_fam; 2.
 DR Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
 KW Transit peptide.
 FT TRANSIT 1 69 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 70 448 OMEGA-6 FATTY ACID DESATURASE.
 FT DOMAIN 171 175 HISTIDINE BOX-1.
 FT DOMAIN 207 211 HISTIDINE BOX-2.
 FT DOMAIN 367 371 HISTIDINE BOX-3.
 SQ SEQUENCE 448 AA; 51225 MW; C3AC72FB28FBF287 CRC64;
 Query Match 89.5%; Score 34; DB 1; Length 448;

Best Local Similarity 66.7%; Pred. No. 38;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 257 WVNWHF 262
 RESULT 8
 CATA_PSEPU STANDARD; PRT; 479 AA.
 ID 059714;
 AC 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Catalase (EC 1.11.1.6).
 GN KATA OR CATA.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Corvallis;
 RX MEDLINE=98019091; PubMed=9358059;
 RA Kim Y.C., Miller C.D., Anderson A.J.;
 RA "Identification of adjacent genes encoding the major catalase and a
 RT bacterioferritin from the plant-beneficial bacterium Pseudomonas
 putida";
 RL Gene 199:219-224 (1997).
 CC -!- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;
 CC serves to protect cells from the toxic effects of hydrogen
 CC peroxide.
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -!- COFACTOR: Heme group.
 CC -!- ENZYME REGULATION: ACTIVATED BY PEROXIDE.
 CC -!- SIMILARITY: Belongs to the catalase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U63511; AAB88219.1; -;
 DR HSPSP; P42321; 2CAE.
 DR InterPro; IPR002226; Catalase.
 DR Pfam; PF00199; catalase; 1.
 DR PRINTS; PR00067; CATALASE.
 DR ProDom; PD000510; Catalase; 1.
 DR PROSITE; PS00437; CATALASE_1; 1.
 DR PROSITE; PS00438; CATALASE_2; 1.
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
 KW ACT SITE 53 53 BY SIMILARITY.
 FT ACT SITE 126 126 BY SIMILARITY.
 FT METAL 336 336 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 479 AA; 53381 MW; EFECBDE67778571 CRC64;
 Query Match 89.5%; Score 34; DB 1; Length 479;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 209 WVKWHF 214
 RESULT 9
 SIK1_YEAST STANDARD; PRT; 504 AA.
 ID SIK1_YEAST
 AC Q12460;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE SIK1 protein (Nucleolar protein NOP56).
 GN SIK1 OR NOP56 OR YLR197W OR L8167.9.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / YPH1;
 RX MEDLINE=96040178; PubMed=7547500;
 RA Morin P.J., Downs J.A., Snodgrass A.M., Gilmore T.D.;
 RT "Genetic analysis of growth inhibition by GAL4-L kappa B-alpha in
 RT Saccharomyces cerevisiae.";
 RL Cell Growth Differ. 6:789-798 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Ansoerge W.,
 RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
 RA Enian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 RA Heuss-Neitzel D., Halbert H., Hilger F., Kleine K., Koetter P.,
 RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
 RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
 RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
 RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhaesselt P.,
 RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
 RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hohelsel J.D.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 RL Nature 387:87-90 (1997).
 RN [3]
 RP CHARACTERIZATION, AND MUTAGENESIS.
 RC MEDLINE=98038777; PubMed=9372940;
 RX Gautier T., Berges T., Tollervey D., Hurt E.;
 RA "Nucleolar KXE/D repeat proteins Nop56p and Nop58p interact with Nop1p
 RT and are required for ribosome biogenesis.";
 RL Mol. Cell. Biol. 17:7088-7098 (1997).
 CC [1]- FUNCTION: Required for 60S ribosomal subunit synthesis.
 CC [1]- SUBUNIT: Interacts with Nop1 and Nop58.
 CC [1]- SUBCELLULAR LOCATION: Nuclear; nucleolar.
 CC [1]- SIMILARITY: Belongs to the NOP5/NOP56 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U20237; AAC49066.1; --
 CC EMBL; U14913; AAB67431.1; --
 CC PIR; S48550; S48550.
 CC Germonline; 142259; --
 CC SGD; S0004187; SIK1.
 CC GO; GO:0005732; C:small nucleolar ribonucleoprotein complex; IPI.
 CC GO; GO:0030490; P:processing of 20S pre-rRNA; IPI.
 CC InterPro; IPR002687; Nop.
 CC Pfam; PF01798; Nop; 1.
 CC ProDom; PD004104; Nop; 1.
 CC DOMAIN 443 504
 CC MUTAGEN 333 333
 CC MUTAGEN 355 355
 CC V-S-A: REDUCED GROWTH RATE AT ALL
 CC TEMPERATURES; WHEN ASSOCIATED WITH R-385.
 CC Y-S-C: AT 37 DEGREES, GROWTH SLOWS AFTER 6
 CC TO 8 HOURS AND CELL DIVISION STOPS AFTER
 CC 20 HOURS.
 CC M-R-R: REDUCED GROWTH RATE AT ALL
 CC TEMPERATURES; WHEN ASSOCIATED WITH A-333.

SQ SEQUENCE 504 AA; 56864 MW; F8522A5870EF4842 CRC64;
 Query Match 89.5%; Score 34; DB 1; Length 504;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 DB 198 WYGHF 203
 RESULT 10
 NOP5_YEAST
 ID NOP5_YEAST STANDARD; PRT; 511 AA.
 AC Q12499;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nucleolar protein NOP58 (Nucleolar protein NOP5).
 GN NOP58 OR NOP5 OR YOR310C OR O6108.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / FY1679;
 RA Pearson B.M., Hernando Y., Wolf S.S., Kalogeropoulos A., Schweizer M.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=W303;
 RX MEDLINE=98298165; PubMed=9632712;
 RA Wu P., Brockenbrough J.S., Metcalfe A.C., Chen S., Aris J.P.;
 RT "Nop5p is a small nucleolar ribonucleoprotein component required for
 RT pre-18S rRNA processing in yeast.";
 RL Biol. Chem. 273:16453-16463 (1998).
 CC [1]- FUNCTION: Required for pre-18S rRNA processing. May bind
 CC microtubules.
 CC [1]- SUBUNIT: Interacts with NOP56 and NOP1.
 CC [1]- SUBCELLULAR LOCATION: Nuclear; nucleolar.
 CC [1]- SIMILARITY: Belongs to the NOP5/NOP56 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X90565; CAA62165.1; --
 CC EMBL; Z75217; CAA99630.1; --
 CC PIR; S58322; S58322.
 CC Germonline; 143898; --
 CC SGD; S0005837; NOP58.
 CC GO; GO:0005732; C:small nucleolar ribonucleoprotein complex; IPI.
 CC GO; GO:0003754; F:chaperone activity; NAS.
 CC GO; GO:0017069; F:snRNA binding; IDA.
 CC GO; GO:0030490; P:processing of 20S pre-rRNA; IPI.
 CC GO; GO:0006608; P:snRNP protein-nucleus import; NAS.
 CC InterPro; IPR002687; Nop.
 CC Pfam; PF01798; Nop; 1.
 CC ProDom; PD004104; Nop; 1.
 CC Y-S-C: AT 37 DEGREES, GROWTH SLOWS AFTER 6
 CC TO 8 HOURS AND CELL DIVISION STOPS AFTER
 CC 20 HOURS.
 CC M-R-R: REDUCED GROWTH RATE AT ALL
 CC TEMPERATURES; WHEN ASSOCIATED WITH A-333.

SQ SEQUENCE 511 AA; 56956 MW; 8A28894482A19E2 CRC64;
 Query Match 89.5%; Score 34; DB 1; Length 511;
 Best Local Similarity 66.7%; Pred. No. 43;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;


```
QY 1 WXXWHF 6
Db 187 WYGMHF 192

RESULT 11
NOP5 HUMAN
ID NOP5 HUMAN STANDARD; PRT; 529 AA.
AC Q9Y2X3; Q9P036; Q9UFN3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nucleolar protein NOP5 (Nucleolar protein 5) (NOP58) (HSPC120).
GN NOP5
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lyman S.K., Grace L.;
RT "Cloning and characterization of NOP5/NOP58.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Nelson S.A., Santora K.E., LaRoche W.J.;
RT "Isolation and characterization of a novel PDGF-induced human gene.";
RL Gene 253:87-93 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.J., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.D., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 1-447 FROM N.A.
RA Tissue=Brain;
RA Bloeker H., Boecher M., Brandt P., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 64-529 FROM N.A.
RA Tissue=Blood;
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;
RT "Cloning and functional analysis of cDNAs with open reading frames for
300 previously undefined genes expressed in CD34+ hematopoietic
stem/progenitor cells.";
RL Genome Res. 10:1546-1560 (2000).
CC -!- FUNCTION: Required for 60S ribosomal subunit biogenesis (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.

CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Belongs to the NOP5/NOP56 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF123534; AAD27610.1; -
CC EMBL; AF263608; AAF91394.1; -
CC EMBL; BC032592; AAB32592.1; -
CC EMBL; AL117554; CAB55989.1; -
CC EMBL; AF161469; AAF29084.1; -
CC PIR; T17299; T17299.
CC SWISS-2DPAGE; Q9Y2X3; HUMAN.
CC GO; GO:0005730; C:nucleolus; TAS.
CC GO; GO:0003754; F:chaperone activity; ISS.
CC GO; GO:00030519; F:snRNP binding; ISS.
CC GO; GO:0016049; P:cell growth; TAS.
CC GO; GO:0006364; P:rRNA processing; TAS.
CC GO; GO:0006608; P:snRNP protein-nucleus import; ISS.
CC InterPro; IPR002887; Nop.
CC Pfam; PF01798; Nop; 1.
CC ProDom; PD004104; Nop; 1.
CC Ribosome biogenesis; Nuclear protein.
CC CONFLICT 2 129 G -> M (IN REF. 4).
CC CONFLICT 202 221 G -> V (IN REF. 4).
CC CONFLICT 235 260 LTYCKLQKVGDRKNYASAK -> YHTASVTRKLAIGRLCL
CC CONFLICT 280 280 KAAAEISMGTSVEEDICNIHLCTQ -> EGSCRDHNGNR
CC CONFLICT 443 447 GFRRLYQYASLHP (IN REF. 5).
CC CONFLICT 447 447 M -> V (IN REF. 4).
CC SEQUENCE 529 AA; 59578 MW; 27CD73CF5B9A556 CRC64;
Query Match 89.5%; Score 34; DB 1; Length 529;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 WXXWHF 6
Db 186 WYGMHF 191
RESULT 12
NOP5 RAT
ID NOP5 RAT STANDARD; PRT; 534 AA.
AC Q9QZ86; O88525;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nucleolar protein NOP5 (Nucleolar protein 5) (Noppl40 associated
protein).
DE protein.
GN NOP5 OR NAP65.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 1-22, AND INTERACTION WITH NOLC1.
RX MEDLINE=20143579; PubMed=10679015;
RA Yang Y., Isaac C., Wang C., Dragon F., Pogacic V., Meier U.T.;
RT "Conserved composition of mammalian box H/ACA and box C/D small
nucleolar ribonucleoprotein particles and their interaction with the
common factor Noppl40."
RL Mol. Biol. Cell 11:567-577 (2000).
RN [2]
RP SEQUENCE OF 1-461 FROM N.A.
RA Hattori D., Gray J.C.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
```

```

CC CC -!- FUNCTION: Required for 60S ribosomal subunit biogenesis (By
CC CC similarity).
CC CC -!- SUBUNIT: Interacts with Nop1/Nop140.
CC CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC CC -!- SIMILARITY: Belongs to the NOP5/NOP56 family.
CC CC
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; AF194371; AAF05769.1; -
CC CC EMBL; AF069782; AAC23535.1; -
CC CC DR GO; GO:0005730; C:nucleolus; ISS.
CC CC DR GO; GO:0003754; F:chaperone activity; TAS.
CC CC DR GO; GO:00030519; F:snRNP binding; TAS.
CC CC DR GO; GO:0016049; F:cell growth; ISS.
CC CC DR GO; GO:0006364; F:rRNA processing; ISS.
CC CC DR GO; GO:0006608; F:snRNP protein-nucleus import; IDA.
CC CC DR InterPro: IPR002687; Nop.
CC CC DR Pfam: PF01798; Nop; 1.
CC CC DR ProDom: PD004104; Nop; 1.
CC CC KW Ribosome biogenesis; Nuclear protein.
CC CC FT CONFLICT 396 396 R -> K (IN REF. 2).
CC CC FT CONFLICT 459 459 A -> K (IN REF. 2).
CC CC SQ SEQUENCE 534 AA; 60070 MW; 4B9585FA14E67799 CRC64;
CC CC
CC CC Query Match 89.5%; Score 34; DB 1; Length 534;
CC CC Best Local Similarity 66.7%; Pred. No. 45;
CC CC Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CC CC
CC CC QY 1 WXXWHF 6
CC CC Db 186 WYGWHF 191
CC CC
CC CC RESULT 13
CC CC LYSA DROME
CC CC ID LYSA DROME STANDARD; .PRT; 140 AA.
CC CC AC P37157; P29614; Q9W0J6; Q9W0J7;
CC CC DT 01-OCT-1994 (Rel. 30, Created)
CC CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC CC DE Lysozyme A/C/D precursor (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase
CC CC A/C).
CC CC GN LYSA AND (LYSC OR CG9111) AND (LYSD OR CG9118).
CC CC OS Drosophila melanogaster (Fruit fly).
CC CC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC CC OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC CC OC Ephydroidea; Drosophilidae; Drosophila.
CC CC OX NCBI_TaxID=7227;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RC STRAIN=Canton-S;
CC CC RX MEDLINE=94211204; PubMed=8159165;
CC CC RA Daffre S., Kystlen P., Samakovlis C., Hultmark D.;
CC CC RT "The lysozyme locus in Drosophila melanogaster: an expanded gene
CC CC family adapted for expression in the digestive tract.";
CC CC RL Mol. Gen. Genet. 242:152-162(1994).
CC CC RN [2]
CC CC RP SEQUENCE FROM N.A.
CC CC RC STRAIN=Berkley;
CC CC RX MEDLINE=20196006; PubMed=10731132;
CC CC RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
CC CC RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
CC CC RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
CC CC RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
CC CC RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
CC CC RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
CC CC RA Abril J.F., Agbayani A., An H.-J., Andrews-Ffankoch C., Baldwin D.,

```

```

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Buschan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
RL
CC CC -!- FUNCTION: Unlikely to play an active role in the humoral immune
CC CC defense. May have a function in the digestion of bacteria in the
CC CC food.
CC CC
CC CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
CC CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
CC CC heteropolymers of the prokaryotes cell walls.
CC CC
CC CC -!- TISSUE SPECIFICITY: Found in the midgut.
CC CC
CC CC -!- DEVELOPMENTAL STAGE: Maximal expression is found during the third
CC CC larval instar, it drops to become undetectable in the late pupal
CC CC stage. The expression in adults is similar to that of first and
CC CC second larval instars.
CC CC
CC CC -!- SIMILARITY: Belongs to family 22 of glycosyl hydrolases.
CC CC
CC CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC CC gene model prediction.
CC CC
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; Z22223; CAA80225.1; -
CC CC EMBL; Z22226; CAA80228.1; -
CC CC DR EMBL; X58382; CAA1272.1; -
CC CC DR EMBL; AE003470; AAF47449.1; ALT_SEQ.
CC CC EMBL; AE003470; AAF47450.1; -
CC CC DR EMBL; AE003470; -; NOT_ANNOTATED_CDS.
CC CC PIR; S20914; S20914.
CC CC PIR; S41573; S41573.
CC CC HSSP; P00695; ILZ5.
CC CC FlyBase; FBgn0011201; Lysa.
CC CC FlyBase; FBgn0004426; LysC.
CC CC FlyBase; FBgn0004427; LysD.
CC CC GO; GO:0004568; F:chitinase activity; IDA.
CC CC GO; GO:0003796; F:lysozyme activity; IDA.
CC CC InterPro: IPR001916; Glyco_hydro_22.
CC CC Pfam; PF00062; lys; 1.
CC CC PRINTS; PR00135; LYZLACT.
CC CC SMART; SM00263; LY21; 1.
CC CC PROSITE; PS00128; LACTALBUMIN_LYSOZYME; 1.

```

DT	02-OCT-2004	(Rel. 30, Created)
DT	16-OCT-2001	(Rel. 40, Last sequence update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)

Lysozyme E precursor (EC 3.2.1.17) (1,4-beta-N-acetyl-muramidase E).
 LYSE OR CG180.
 Drosophila melanogaster (Fruit fly).
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=Canton-S;
 RX MEDLINE=94211204; PubMed=8159165;
 RA Daffre S., Kystlen P., Samakovilis C., Hultmark D.;
 RT "The lysozyme locus in *Drosophila melanogaster*: an expanded gene
 family adapted for expression in the digestive tract.";
 RL Mol. Gen. Genet. 242:152-162(1994).
 [2]
 SEQUENCE FROM N.A.
 STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Abril J.F., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Balow R.M., Basu A., Baxendale J., Andrews-Pfannkuch C., Baldwin D.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brackstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*";
 RL Science 287:2185-2195(2000).
 CC -!- FUNCTION: Unlikely to play an active role in the humoral immune
 CC defense. May have a function in the digestion of bacteria in the
 CC food.
 CC
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
 CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
 CC heteropolymers of the prokaryotes cell walls.
 CC
 CC -!- TISSUE SPECIFICITY: Found in the midgut.
 CC
 CC -!- DEVELOPMENTAL STAGE: Maximal expression is found during the third
 CC larval instar, it drops to become undetectable in the late pupal
 CC stage. The expression in adults is similar to that of first and
 CC second larval instars.
 CC
 CC -!- SIMILARITY: Belongs to family 22 of glycosyl hydrolases.
 CC
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.

 CC EMBL; Z22227; CAA80229.1; -;
 CC EMBL; AE003470; AAF47451.1; -;
 CC PIR; S41577; S41577.
 CC HSP; P00695; LOUE.
 CC FlyBase; FBgn0004428; Lyse.
 CC InterPro; IPR001916; Glyco_hydro_22.
 CC Pfam; PF00062; lys_1
 CC PRINTS; PR00135; LYSLACT.
 CC SMART; SM00263; LY21; 1.
 CC PROSITE; PS00128; LACTALBUMIN LYSOZYME; 1.
 CC Hydrolase; Glycosidase; Bacteriolytic enzyme; Signal;
 CC Multigene family.
 CC SIGNAL 1 18 BY SIMILARITY.
 CC CHAIN 19 140 LYSOZYME E.
 CC DISULFID 24 139 BY SIMILARITY.
 CC DISULFID 45 129 BY SIMILARITY.
 CC DISULFID 80 96 BY SIMILARITY.
 CC DISULFID 92 110 BY SIMILARITY.
 CC ACT_SITE 50 50 BY SIMILARITY.
 CC ACT_SITE 68 68 BY SIMILARITY.
 CC CONFLICT 12 12 M -> L (IN REF. 1).
 CC CONFLICT 76 76 N -> D (IN REF. 1).
 CC CONFLICT 138 138 G -> D (IN REF. 1).
 CC SEQUENCE 140 AA; 15552 MW; CEB5465CF6B6F123 CRC64;
 Query Match 86.8%; Score 33; DB 1; Length 140;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 123 WSTWHY 128

 RESULT 16
 LYSE DROME
 ID LYSE DROME STANDARD; PRT; 142 AA.
 AC P37161; O9WOK1;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 10-OCT-2003 (Rel. 42, last annotation update)
 DE Lysozyme X precursor (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase X).
 GN LYSE OR CG9120.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RX SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Abril J.F., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Balow R.M., Basu A., Baxendale J., Andrews-Pfannkuch C., Baldwin D.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brackstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasako P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimias I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE OF 62-142 FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=94211204; PubMed=8159165;
RA Daffre S., Kylsten P., Samakowis C., Hultmark D.;
RT "The lysozyme locus in *Drosophila melanogaster*: an expanded gene
family adapted for expression in the digestive tract.";
RL Mol. Gen. Genet. 242:152-162 (1994).
CC -!- FUNCTION: UNLIKELY TO PLAY AN ACTIVE ROLE IN THE HUMORAL IMMUNE
DEFENSE. MAY HAVE A FUNCTION IN THE DIGESTION OF BACTERIA IN THE
FOOD. MAY BE INVOLVED IN THE CLEARANCE OF BACTERIA FROM THE LARVAL
GUT BEFORE METAMORPHOSIS.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
heteropolymers of the prokaryotes cell walls.
CC -!- TISSUE SPECIFICITY: Found in the midgut.
CC -!- DEVELOPMENTAL STAGE: RISES DRAMATICALLY IN THE LATE THIRD INSTAR,
THEN DECREASES GRADUALLY DURING THE PUPAL STAGES. LOW EXPRESSION
IS FOUND IN ADULTS.
CC -!- SIMILARITY: Belongs to family 22 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
DR EMBL; AE003470; AAF47445.1; --
DR EMBL; Z22224; CA80226.1; --
DR PIR; S41580; S41580.
DR HSSP; P00698; IAT5.
DR FlyBase; FBgn004431; LvsX.
DR InterPro; IPR001916; Glyco_hydro_22.
DR Pfam; PF00062; lvs; 1.
DR PRINTS; PR00135; LYZLACT.
DR SMART; SM00263; LYZ1; 1.
DR PROSITE; PS00128; LACTALBUMIN LYSOZYME; 1.
KW Hydrolase; Glycosidase; Bacteriolytic enzyme; Signal;
KW Multigene family.
FT SIGNAL 1 19
FT CHAIN 20 142
FT DISULFID 25 140
FT DISULFID 46 130
FT DISULFID 81 97
FT DISULFID 93 111
FT ACT_SITE 51 51
FT ACT_SITE 69 69
FT ACT_SITE 78 78
FT CONFLICT 78 M -> L (IN REF. 2).
SQ SEQUENCE 142 AA; 15591 MW; 2A48035364B995BC CRC64;

Query Match 86.8%; Score 33; DB 1; Length 142;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 124 WSTWHY 129

RESULT 17

Y4CD RHISN
ID Y4CD RHISN STANDARD; PRT; 640 AA.
AC P55386;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 71.6 kDa protein Y4CD.
GN Y4CD.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401 (1997).
CC -!- SIMILARITY: None obvious.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
DR EMBL; AE000067; AAB91634.1; --
DR PIR; T28631; T28631.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 640 AA; 71609 MW; 829BF90C595A3C0B CRC64;

Query Match 86.8%; Score 33; DB 1; Length 640;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 521 WPFWHF 526

RESULT 18

PLSB CAEEL
ID PLSB CAEEL STANDARD; PRT; 718 AA.
AC Q22949;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Probable glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE (EC 2.3.1.15) (GPAT).
GN F08F3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Blanchard M., Bradshaw H.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
 CC acyl-glycerol 3-phosphate.
 CC -1- PATHWAY: First step in de novo phospholipid biosynthesis. It may
 CC also function in the regulation of membrane biogenesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC (Potential).
 CC -1- SIMILARITY: Belongs to the GPAT / DAPAT family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U64847; AAB04876.1; -;
 DR PIR; T29448; T29448.
 DR WormPep; F08F3.2; CE09258.
 DR InterPro; IPR001213; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.
 DR SMART; SM00563; Pfam; 1.
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;
 KW Transmembrane; Mitochondrion; Transmembrane; TRANSIT ?
 FT CHAIN 1 ?
 FT TRANSIT ?
 FT CHAIN 1 ?
 FT TRANSIT ?
 FT TRANSMEM 409 425
 FT SEQUENCE 718 AA; 82071 MW; E0A36A4A86FC138D CRC64;
 SQ
 Query Match 86.8%; Score 33; DB 1; Length 718;
 Best Local Similarity 66.7%; Pred No. 87;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 178 WCNWHF 183

 RESULT 19
 LYSS DROME STANDARD; PRT; 140 AA.
 ID LYSS DROME STANDARD; PRT; 140 AA.
 AC P37160; Q9W0U3;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Lysozyme S precursor (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase S).
 GN LYSS OR CG1165.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RC MEDLINE=94211204; PubMed=8159165;
 RA Daffre S., Kylisten P., Samakovits C., Hultmark D.;
 RT "The lysozyme locus in Drosophila melanogaster: an expanded gene
 RT family adapted for expression in the digestive tract.";
 RL Mol. Gen. Genet. 242:152-162(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fiesler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.R., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laake P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: Unlikely to play an active role in the humoral immune
 CC defense. May have a function in the digestion of bacteria in the
 CC food.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
 CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
 CC heteropolymers of the prokaryotes cell walls.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE MIDGUT, ESPECIALLY THE EPITHELIUM
 CC OF THE GASTRIC CAECAE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED ONLY DURING THE LARVAL STAGE, THE
 CC HIGHEST LEVEL IS REACHED DURING THE THIRD LARVAL INSTAR.
 CC -1- SIMILARITY: Belongs to family 22 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE003470; AAF47453.1; -;
 DR EMBL; Z22228; CAA80230.1; -;
 DR PIR; S41579; S41579.
 DR HGSP; P00695; IC46.
 DR FlyBase; FBgn004430; Lyss.
 DR InterPro; IPR01916; Glyco_hydro_22.
 DR PRINTS; PR00062; Lys; 1.
 DR Pfam; PF00062; Lys; 1.
 DR SMART; SM00263; LY21; 1.
 DR PROSITE; PS00128; LACTALBUMIN LYSOZYME; 1.
 DR Hydrolase; Glycosidase; Bacteriolytic enzyme; Signal;
 KW Multigene family.
 KW SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 140
 FT DISULFID 25 139 BY SIMILARITY.
 FT DISULFID 46 129 BY SIMILARITY.
 FT DISULFID 81 96 BY SIMILARITY.
 FT DISULFID 92 110 BY SIMILARITY.
 FT ACT_SITE 51 69 BY SIMILARITY.
 FT ACT_SITE 69 69 BY SIMILARITY.
 FT CONFLICT 11 16 AIAAPA -> PLPLC (IN REF. 1).
 SQ SEQUENCE 140 AA; 15651 MW; ACD139CG656EF8FC CRC64;

Query Match 84.2%; Score 32; DB 1; Length 140;
 Best Local Similarity 50.0%; Pred. No. 28;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
 Db 123 WAVWHY 128

RESULT 20
 YC60 MYCTU STANDARD; PRT; 372 AA.
 ID YC60 MYCTU
 AC Q11058;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein Rv1260/MT1298.
 GN Rv1260 OR MT1298 OR MTCY50.22C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RA Felschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490(2002).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z77137; CAB00893.1; -;
 CC DR EMBL; AE007005; AAK45557.1; ALT_INIT.
 CC DR PIR; D70753; D70753.
 CC DR TIGR; MT1298; -;
 CC DR TubercuList; Rv1260; -;
 CC DR InterPro; IPR000733; Flav_monooxygenase.
 CC DR InterPro; IPR003042; Rng_monooxygenase.
 CC DR Pfam; PF01360; Monooxygenase; 1.
 CC DR PRINTS; PR00420; RNMNOXGNASE.
 CC KW Hypothetical protein; Complete proteome.
 CC SEQUENCE 372 AA; 41329 MW; C2E5BE08D6E29FF CRC64;

Query Match 84.2%; Score 32; DB 1; Length 372;
 Best Local Similarity 50.0%; Pred. No. 70;

Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 WXXWHF 6
 Db 195 WQTHY 200

RESULT 21
 BCR_ECOLI
 ID BCR_ECOLI STANDARD; PRT; 396 AA.
 AC P28246;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bicyclomycin resistance protein (Sulfonamide resistance protein).
 GN BCR OR BICA OR BICR OR SUR OR SUXA OR B2182.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nichols B.P.;
 RL Unpublished observations (FEB-1994).
 RN [2]
 RP SEQUENCE OF 60-396 FROM N.A.
 RC STRAIN=K12 / C600;
 RX MEDLINE=93252267; PubMed=8486276;
 RA Bentley J., Hyatt L.S., Ainley K., Parish J.H., Herbert R.B.,
 RA White G.R.;
 RT "Cloning and sequence analysis of an Escherichia coli gene conferring
 RL bicyclomycin resistance."
 RN Gene 127:117-120(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / BHB2600;
 RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
 RA Church G.M.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1234-1238(1997).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=90146244; PubMed=2694948;
 RA Nichols B.P., Guay G.G.;
 RL "Gene amplification contributes to sulfonamide resistance in
 RT Escherichia coli."
 CC -1- FUNCTION: Involved in sulfonamide (sulfathiazole) and
 CC bicyclomycin resistance. Probable membrane translocase.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY). BELONGS TO THE
 CC BCR/CMLA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X63703; CAA45230.1; ALT_INIT.
 CC DR EMBL; U000008; AAA16406.1; ALT_INIT.

```

DR EMBL; AE000308; AAC75243.1; -.
DR PIR; E64987; E64987.
DR EcoGene; EG11419; bcr.
DR InterPro; IPR004734; Drug resist.
DR InterPro; IPR004812; Efflux_Bcr_CflA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR TIGRFAMs; TIGR00880; 2 A.01.02; 1.
DR TIGRFAMs; TIGR00710; efflux_Bcr_CflA; 1.
DR PROSITE; PS50850; MFS; 1.
DR Antibiotic resistance; Transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 9 26 POTENTIAL.
FT TRANSMEM 49 65 POTENTIAL.
FT TRANSMEM 77 98 POTENTIAL.
FT TRANSMEM 106 122 POTENTIAL.
FT TRANSMEM 138 159 POTENTIAL.
FT TRANSMEM 166 185 POTENTIAL.
FT TRANSMEM 215 239 POTENTIAL.
FT TRANSMEM 251 268 POTENTIAL.
FT TRANSMEM 286 302 POTENTIAL.
FT TRANSMEM 309 332 POTENTIAL.
FT TRANSMEM 345 366 POTENTIAL.
FT TRANSMEM 373 390 POTENTIAL.
FT CONFLICT 99 99 D -> V (IN REF. 2).
FT CONFLICT 246 246 V -> I (IN REF. 1 AND 2).
SQ SEQUENCE 396 AA; 43352 MW; D609AE35370B6AID CRC64;

Query Match 84.2%; Score 32; DB 1; Length 396;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 160 WLSWHY 165

RESULT 22
SHT3_RAT
ID SHT3_RAT STANDARD; PRT; 483 AA.
AC P35563.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5-hydroxytryptamine 3 receptor precursor (5-HT-3) (Serotonin-gated ion
channel receptor) (5-HT3R).
GN HTR3A OR HTR3 OR 5HT3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Miyake A., Mochizuki S., Akuzawa S., Kon G.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 22-483 FROM N.A.
RX MEDLINE=94154206; PubMed=7509203;
RA Iseberg K.E., Ukhun I.A., Holstad S.G., Jafri S., Uchida U.,
RA Zorunski C.F., Yang J.;
RT "Partial cDNA cloning and NGF regulation of a rat 5-HT3 receptor
subunit.";
RL NeuroReport 5:121-124(1993).
CC -!- FUNCTION: This is one of the several different receptors for
5-hydroxytryptamine (serotonin), a biogenic hormone that functions
as a neurotransmitter, a hormone, and a mitogen. This receptor is
a ligand-gated ion channel, which when activated causes fast,
depolarizing responses in neurons. It is a cation-specific, but
otherwise relatively nonselective, ion channel.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; D49395; BAA08388.1; -.
EMBL; U01227; AAA52182.1; -.
DR InterPro; IPR008132; 5HT3_receptor.
DR InterPro; IPR008133; 5HT3_receptor_A.
DR InterPro; IPR006029; Neu_chan_memb.
DR InterPro; IPR006202; Neu_chan_LBD.
DR InterPro; IPR006201; Neu_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR01709; 5HT3ARECEPTR.
DR PRINTS; PR01708; 5HT3RECEPTOR.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Ionic channel; Transmembrane; Receptor; Postsynaptic membrane;
KW Glycoprotein; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 483 5-HYDROXYTRYPTAMINE 3 RECEPTOR.
FT DOMAIN 24 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 273 1 (POTENTIAL).
FT DOMAIN 274 278 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 279 297 2 (POTENTIAL).
FT DOMAIN 298 307 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 308 326 3 (POTENTIAL).
FT DOMAIN 327 460 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 461 480 4 (POTENTIAL).
FT DOMAIN 481 483 EXTRACELLULAR (POTENTIAL).
FT DISULFID 162 176 BY SIMILARITY.
FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 191 191 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 22 22 E -> K (IN REF. 2).
FT CONFLICT 306 306 G -> R (IN REF. 2).
SQ SEQUENCE 483 AA; 55428 MW; ED85257BBCCF28A4 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 483;
Best Local Similarity 50.0%; Pred. No. 89;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
DB 477 WSIWHY 482

RESULT 23
SHT3_MOUSE
ID SHT3_MOUSE STANDARD; PRT; 487 AA.
AC P29979; Q61225; Q61226;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5-hydroxytryptamine 3 receptor precursor (5-HT-3) (Serotonin-gated ion
channel receptor) (5-HT3R).
GN HTR3A OR HTR3 OR 5HT3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92022603; PubMed=1718042;
RA Maricq A.V., Peterson A.S., Brake A.J., Myers R.M., Julius D.;
RT "Primary structure and functional expression of the 5HT3 receptor, a
serotonin-gated ion channel.";

```


Science 254:432-437(1991).

[2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE=Brain;
 RX MEDLINE=94156052; PubMed=8112471;
 RA Uetz P., Abdelatty F., Villarroel A., Gundrun R., Weiss B., Koenen M.;
 RT "Organisation of the murine 5-HT3 receptor gene and assignment to
 human chromosome 11.";
 RL FEBS Lett. 339:302-306(1994).

[3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A/J;
 RX MEDLINE=93259238; PubMed=7683998;
 RA Hope A.G., Downie D.L., Sutherland L., Lambert J.J., Peters J.A.,
 RA Burchell B.;
 RT "Cloning and functional expression of an apparent splice variant of
 the murine 5-HT3 receptor A subunit.";
 RL Eur. J. Pharmacol. 245:187-192(1993).

[4]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC STRAIN=129/Sv;
 RX MEDLINE=95157178; PubMed=7854052;
 RA Werner P., Kawashima E., Reid J., Hussy N., Lundstrom K., Buell G.,
 RA Humbert Y., Jones K.A.;
 RT "Organization of the mouse 5-HT3 receptor gene and functional
 expression of two splice variants.";
 RL Brain Res. Mol. Brain Res. 26:233-241(1994).

-!- FUNCTION: This is one of the several different receptors for 5-
 hydroxytryptamine (serotonin), a biogenic hormone that functions
 as a neurotransmitter, a hormone, and a mitogen. This receptor is
 a ligand-gated ion channel, which when activated causes fast,
 depolarizing responses in neurons. It is a cation-specific, but
 otherwise relatively nonselective, ion channel.
 -!- SUBCELLULAR LOCATION: Integral membrane protein.
 -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Name=5-HT3R-A;
 IsoId=P233979-1; Sequence=Displayed;
 Name=5-HT3R-AS;
 IsoId=P233979-2; Sequence=VSP 000079;
 -!- TISSUE SPECIFICITY: Brain, spinal cord, and heart.
 -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; M74425; AAA37124.1; -;
 DR EMBL; Z22772; CAA80453.1; -;
 DR EMBL; Z22773; CAA80453.1; JOINED.
 DR EMBL; X72395; CAA51089.1; -;
 DR EMBL; X79283; CAA55870.1; -;
 DR EMBL; X79283; CAA55871.1; -;
 DR PIR; S41757; S41757.
 DR MGD; MGI:96282; Htr3a.
 DR InterPro; IPR008132; 5HT3 receptor.
 DR InterPro; IPR008133; 5HT3_receptor_A.
 DR InterPro; IPR006023; Neu_chan_memb.
 DR InterPro; IPR006202; Neur_chan_LBD.
 DR InterPro; IPR006201; Neur channel.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR01709; 5HT3ARECPT.
 DR PRINTS; PR01708; 5HT3RECEPT.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFS; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 DR Ionic channel; Transmembrane Receptor; Postsynaptic membrane;
 Glycoprotein; Signal; Alternative splicing.

FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 487 5-HYDROXYTRYPTAMINE 3 RECEPTOR.
 FT DOMAIN 24 245 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 246 272 1 (POTENTIAL).
 FT DOMAIN 273 277 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 278 296 2 (POTENTIAL).
 FT DOMAIN 297 305 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 306 324 3 (POTENTIAL).
 FT DOMAIN 325 464 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 465 484 4 (POTENTIAL).
 FT DOMAIN 485 487 EXTRACELLULAR (POTENTIAL).
 FT DISULFID 161 175 BY SIMILARITY.
 FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 383 388 Missing (in isoform 5-HT3R-AS).
 FT CONFLICT 31 31 /FTId=VSP 000079.
 FT CONFLICT 74 74 E -> AR (IN REF. 3 AND 4).
 FT CONFLICT 302 302 I -> V (IN REF. 3).
 FT CONFLICT 384 384 T -> TA (IN REF. 3 AND 4).
 FT CONFLICT 393 393 S -> SS (IN REF. 2).
 FT CONFLICT 487 487 H -> T (IN REF. 4).
 SQ SEQUENCE 487 AA; 56056 MW; D0148867C8536D66 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 487;
 Best Local Similarity 50.0%; Pred. No. 89;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
 Db 481 WSIWHY 486

RESULT 24
 5HT3 CAVPO
 ID 5HT3 CAVPO STANDARD; PRT; 490 AA.
 AC 070212; 070213;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 5-hydroxytryptamine 3 receptor precursor (5-HT-3) (Serotonin-gated ion
 channel receptor) (5-HT3R).
 GN HTR3A OR HTR3 OR 5HT3R.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A., (ISOFORMS 1 AND 2), FUNCTION, TISSUE
 SPECIFICITY, AND PHARMACOLOGICAL CHARACTERIZATION.
 RC TISSUE=Small intestine;
 RX MEDLINE=98130690; PubMed=9463477;
 RA Lankiewicz S., Lobitz N., Wetzel C.H.R., Rupprecht R., Gisselmann G.,
 RA Hatt H.;
 RT "Molecular cloning, functional expression, and pharmacological
 characterization of 5-hydroxytryptamine3 receptor cDNA and its splice
 variants from guinea pig.";
 RL Mol. Pharmacol. 53:202-212(1998).

-!- FUNCTION: This is one of the several different receptors for 5-
 hydroxytryptamine (serotonin), a biogenic hormone that functions
 as a neurotransmitter, a hormone, and a mitogen. This receptor is
 a ligand-gated ion channel, which when activated causes fast,
 depolarizing responses in neurons. It is a cation-specific, but
 otherwise relatively nonselective, ion channel.
 -!- SUBCELLULAR LOCATION: Integral membrane protein.
 -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Name=1; Synonyms=5-HT3R-L;
 IsoId=O70212-1; Sequence=Displayed;
 Name=2; Synonyms=5-HT3R-S;
 IsoId=O70212-2; Sequence=VSP 000077;
 -!- TISSUE SPECIFICITY: Expressed in cortex, intestine and liver. Not
 expressed in muscle or spleen.


```

OX NCBI_TaxID=77551;
RN [1]
RN SEQUENCE FROM N.A.
RA Henkle-Duehrsen K.J.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA Henkle-Duehrsen K.J., Eckelt V.H.O., Wildenburg G., Blaxter M.,
RT "Gene structure, activity and localization of a catalase from
RT intracellular bacteria in Onchocerca volvulus.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;
CC serves to protect cells from the toxic effects of hydrogen
CC peroxide.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: Heme group.
CC -!- SIMILARITY: Belongs to the catalase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X82176; CAAS7666.1; -.
DR EMBL; AF069070; AAC79431.1; -.
DR PIR; S49465; S49465.
DR HSSP; P42321; 2CAE.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE 1; 1.
DR PROSITE; PS00438; CATALASE 2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
FT ACT_SITE 55 55 BY SIMILARITY.
FT ACT_SITE 128 128
FT METAL 338 338 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 482 AA; 53700 MW; B1DA5F49BA39F776 CRC64;

Query Match      81.6%; Score 31; DB 1; Length 482;
Best Local Similarity 50.0%; Pred No. 1.3e+02;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 211 WVKWHY 216

RESULT 27
T10B MOUSE
ID T10B MOUSE STANDARD; PRT; 381 AA.
AC Q9Q2M4; Q9JUL5; Q9JUL6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 10B precursor (Death
DE receptor 5) (MK).
GN TNFRSF10B OR DR5 OR KILLER.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=99310501; PubMed=10383128;
RA Wu G.S., Burns T.F., Zhan Y., Alnemri E.S., El-Deiry W.S.;
RT "Molecular cloning and functional analysis of the mouse homologue of
RT the KILLER/DR5 tumor necrosis factor-related apoptosis-inducing ligand

```

```

RT (TRAIL) death receptor.";
RL Cancer Res. 59:2770-2775 (1999).
RN [2]
RN SEQUENCE FROM N.A.
RA Nakamura Y., Tamari M., Watanabe O.;
RL "Mouse TRAIL receptor.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for the cytotoxic ligand TNFSF10/TRAIL. The
CC adaptor molecule FADD recruits caspase-8 to the activated
CC receptor. The resulting death-inducing signaling complex (DISC)
CC performs caspase-8 proteolytic activation which initiates the
CC subsequent cascade of caspases (aspartate-specific cysteine
CC proteases) mediating apoptosis. Promotes the activation of NF-
CC kappaB.
CC -!- SUBUNIT: Homotrimer. Can interact with TRADD and RIP (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in heart, lung and kidney.
CC -!- INDUCTION: TNFSF10B is regulated by the tumor suppressor p53.
CC -!- SIMILARITY: Contains 3 TNFR-Cys repeats.
CC -!- SIMILARITY: Contains 1 death domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF176833; AAD52656.1; -.
DR EMBL; AB031081; BAA96462.1; -.
DR EMBL; AB031082; BAA96463.1; -.
DR HSSP; O14763; 1D0G.
DR MGD; MGI:1341090; Tnfrsf10b.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00020; TNFR_c6; 2.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00208; TNFR; 2.
DR PROSITE; PS00017; DEATH DOMAIN; 1.
DR PROSITE; PS00652; TNFR_NGFR 1; 2.
DR PROSITE; PS00050; TNFR_NGFR 2; 2.
KW Receptor; Apoptosis; Transmembrane; Repeat; Signal.
FT SIGNAL 1 52 POTENTIAL.
FT CHAIN 53 381 TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 10B.
FT DOMAIN 53 160 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 161 181 POTENTIAL.
FT DOMAIN 182 381 CYTOPLASMIC (POTENTIAL).
FT REPEAT 26 86 TNFR-CYS 1.
FT REPEAT 87 129 TNFR-CYS 2.
FT REPEAT 130 169 TNFR-CYS 3.
FT DOMAIN 273 356 DEATH.
FT DISULFID 74 85 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 108 121 BY SIMILARITY.
FT DISULFID 111 129 BY SIMILARITY.
FT DISULFID 131 145 BY SIMILARITY.
FT DISULFID 148 161 BY SIMILARITY.
FT DISULFID 151 169 BY SIMILARITY.
FT CONFLICT 42 42 M -> V (IN REF. 1).
FT CONFLICT 97 97 H -> R (IN REF. 2; BAA96463).
FT CONFLICT 128 128 V -> E (IN REF. 2; BAA96463).
FT CONFLICT 180 180 K -> N (IN REF. 2; BAA96463).
FT CONFLICT 187 187 L -> AT (IN REF. 2; BAA96463).
FT CONFLICT 215 215 R -> RAYP (IN REF. 2; BAA96463).
FT CONFLICT 229 229 L -> S (IN REF. 1).
FT CONFLICT 306 306 K -> R (IN REF. 2; BAA96463).
FT CONFLICT 381 AA; 42223 MW; 2F7999EA2EB485B1 CRC64;
SQ SEQUENCE

```

```
Query Match          78.9%; Score 30; DB 1; Length 381;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 175 WASWH 179

RESULT 28
FIBG PETMA STANDARD; PRT; 432 AA.
ID FIBG PETMA STANDARD; PRT; 554 AA.
AC P04115;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen gamma chain precursor.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=85199776; PubMed=2581603;
RA Strong D.D., Moore M., Cottrell B.A., Bohonus V.L., Pontes M.,
RA Evans B., Riley M., Doolittle R.F.;
RT "Lamprey fibrinogen gamma chain: cloning, cDNA sequencing, and
RT general characterization.";
RL Biochemistry 24:92-101(1985).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot. The soft clot is converted
CC into the hard clot by factor XIIIa which catalyzes the epsilon-
CC (gamma-glutamyl)lysine cross-linking between gamma chains
CC (stronger) and between alpha chains (weaker) of different
CC monomers.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K03049; AAA49262.1; -.
CC PIR; A03129; FGLMGS.
CC DR PDB; 1LWU; 23-AUG-02.
CC DR PDB; 1N73; 07-JAN-03.
CC DR InterPro; IPR002181; Fibrinogen C.
CC DR Pfam; PF00147; fibrinogen_C; 1.
CC DR SMART; SM00186; FBG; 1.
CC DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
CC KW Blood coagulation; Plasma; Glycoprotein; Signal; 3D-structure.
FT CHAIN 1 24 BY SIMILARITY.
FT SINAL 25 432 FIBRINOGEN GAMMA CHAIN.
FT FT 32 32 INTERCHAIN (WITH GAMMA CHAIN)
FT DISULFID 32 32 (BY SIMILARITY).
FT FT 42 42 INTERCHAIN (WITH BETA CHAIN)
FT DISULFID 42 42 (BY SIMILARITY).
FT FT 46 46 INTERCHAIN (WITH ALPHA CHAIN)
FT DISULFID 46 46 (BY SIMILARITY).
FT FT 158 158 INTERCHAIN (WITH BETA CHAIN)
FT DISULFID 158 158 (BY SIMILARITY).
FT FT 162 162 INTERCHAIN (WITH GAMMA CHAIN)
FT DISULFID 162 162 (BY SIMILARITY).
FT FT

Query Match          78.9%; Score 30; DB 1; Length 554;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 391 WATWH 395

RESULT 29
PYRG CORGL STANDARD; PRT; 554 AA.
ID PYRG CORGL STANDARD; PRT; 554 AA.
AC Q8NQL7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CTP synthase (EC 6.3.4.2) (UTP--ammonia ligase) (CTP synthetase).
GN PYRG OR CGL1417.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the ATP-dependent amination of UTP to CTP with
CC either L-glutamine or ammonia as the source of nitrogen (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + UTP + NH(3) = ADP + phosphate + CTP.
CC -!- ENZYME REGULATION: Allosterically activated by GTP, when glutamine
CC is the substrate. Inhibited by CTP (By similarity).
CC -!- PATHWAY: Pyrimidine biosynthesis; conversion of UMP to CTP; third
CC (last) step.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SIMILARITY: Belongs to the CTP synthase family.
CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005278; BAB98810.1; -.
CC DR HAMAP; MF 01227; -.
CC DR InterPro; IPR000991; GATase_1.
CC DR InterPro; IPR004468; Pyrg_synth.
CC DR Pfam; PF00117; GATase; 1.
CC DR TIGRfam; TIGR00337; Pyrg; 1.
CC DR PROSITE; PS00442; GATASE_TYPE_1; FALSE_NEG.
CC KW Pyrimidine biosynthesis; Ligase; Glutamine amidotransferase;
CC Complete proteome.
FT DOMAIN 1 313 AMINATOR DOMAIN.
FT DOMAIN 314 554 GLUTAMINE AMIDOTRANSFERASE.
FT ACT_SITE 394 394 GATASE (BY SIMILARITY).
FT ACT_SITE 526 526 GATASE (BY SIMILARITY).
FT ACT_SITE 528 528 GATASE (BY SIMILARITY).
SQ SEQUENCE 554 AA; 60573 MW; 5A298798580D61FA CRC64;

Query Match          78.9%; Score 30; DB 1; Length 554;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

FT DISULFID 178 207 BY SIMILARITY.
FT DISULFID 348 361 BY SIMILARITY.
FT CARBOHYD 227 N-LINKED (GLUCNAC...).
SQ SEQUENCE 432 AA; 49203 MW; B503979B26DFB24 CRC64;

Query Match          78.9%; Score 30; DB 1; Length 432;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 391 WATWH 395
```

QY 1 WXXWH 5
Db 286 WSTWH 290

RESULT 30
RPOL BPT7
ID -RPOL BPT7 STANDARD; PRT; 883 AA.
AC P00573; Q38543;
DT 21-JUL-1986 (Rel. 01, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-directed RNA polymerase (EC 2.7.7.6).
GN 1.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP SEQUENCE FROM N.A. PubMed=6864790;
RX MEDLINE=83241725;
RA Dunn J.J., Studier F.W.;
RA "Complete nucleotide sequence of bacteriophage T7 DNA and the
RA locations of T7 genetic elements.";
RT J. Mol. Biol. 166:477-535(1983).
RN [2]
RP REVISIONS.
RX MEDLINE=84164887; PubMed=6708104;
RA Moffatt B.A., Dunn J.J., Studier F.W.;
RA "Nucleotide sequence of the gene for bacteriophage T7 RNA
RA polymerase.";
RT J. Mol. Biol. 173:265-269(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82078043; PubMed=7310873;
RA Stahl S.J., Zinn K.;
RA "Nucleotide sequence of the cloned gene for bacteriophage T7 RNA
RA polymerase.";
RT J. Mol. Biol. 148:481-485(1981).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85046624; PubMed=6093820;
RA Grachev M.A., Pletnev A.G.;
RA "Phage T7 RNA-polymerase: gene cloning and its structure.";
RT Biorg. Khim. 10:824-843(1984).
RN [5]
RP SEQUENCE OF 1-59 AND 829-883 FROM N.A.
RX MEDLINE=82078034; PubMed=7310871;
RA Dunn J.J., Studier F.W.;
RA "Nucleotide sequence from the genetic left end of bacteriophage T7
RA DNA to the beginning of gene 4.";
RT J. Mol. Biol. 148:303-330(1981).
RN [6]
RP ACTIVE SITE LYS-631, AND MUTAGENESIS OF LYS-631.
RX MEDLINE=91153326; PubMed=1847871;
RA Maksimova T.G., Mustayev A.A., Zaychikov E.F., Lyakhov D.L.,
RA Tunitskaya V.L., Akbarov A.K., Luchin S.V., Rechinsky V.O.,
RA Chernov B.K., Kochetkov S.N.;
RA "Lys631 residue in the active site of the bacteriophage T7 RNA
RA polymerase. Affinity labeling and site-directed mutagenesis.";
RT Eur. J. Biochem. 195:841-847(1991).
RN [7]
RP MUTAGENESIS OF LYS-172.
RX MEDLINE=93109276; PubMed=1470170;
RA Lyakhov D.L., Ilgenfrits H., Chernov B.K., Dragan S.M.,
RA Rechinsky V.O., Pokholok D.K., Tunitskaya V.L., Kochetkov S.N.;
RA "Site-specific mutagenesis of residue Lys-172 of phage T7 RNA
RA polymerase: characterization of transcription properties of mutant
RA proteins.";
RT Mol. Biol. (Mosk) 26:1022-1035(1992).
RN [8]
RP MUTAGENESIS.
RX MEDLINE=93261429; PubMed=8492813;

RA Rechinsky V.O., Kostyuk D.A., Lyakhov D.L., Chernov B.K.,
RA Kochetkov S.N.;
RT "Random mutagenesis of the gene for bacteriophage T7 RNA polymerase.";
RL Mol. Gen. Genet. 238:455-458(1993).
RN [9]
RP MUTAGENESIS.
RX MEDLINE=93215829; PubMed=8462683;
RA Rechinsky V.O., Tunitskaya V.L., Dragan S.M., Kostyuk D.A.,
RA Kochetkov S.N.;
RT "Tyr-571 is involved in the T7 RNA polymerase binding to its
RT promoter.";
RL FEBS Lett. 320:9-12(1993).
RN [10]
RP MUTAGENESIS.
RX MEDLINE=94190390; PubMed=8133519;
RA Osumi-Davis P., Sreerama N., Volkin D.B., Middaugh C.R., Woody R.W.,
RA Woody A.-Y.M.;
RT "Bacteriophage T7 RNA polymerase and its active-site mutants.
RT Kinetic, spectroscopic and calorimetric characterization.";
RL J. Mol. Biol. 237:5-19(1994).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX MEDLINE=93334440; PubMed=7688864;
RA Sousa R., Chung Y.J., Rose J.P., Wang B.-C.;
RA "Crystal structure of bacteriophage T7 RNA polymerase at 3.3-A
RA resolution.";
RT Nature 364:593-599(1993).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF COMPLEX WITH LYSOZYME.
RX MEDLINE=98336199; PubMed=9670025;
RA Jeruzalmi D., Steitz T.A.;
RA "Structure of T7 RNA polymerase complexed to the transcriptional
RA inhibitor T7 lysozyme.";
RL EMBO J. 17:4101-4113(1998).
RN [13]
RP FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
RP OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
RP SUBSTRATES. RESPONSIBLE FOR THE TRANSCRIPTION OF THE LATE GENES OF
RP T7. IT IS RIFAMPICIN-RESISTANT. IT RECOGNIZES A SPECIFIC PROMOTER
RP SEQUENCE. UNWINDS THE DOUBLE-STRANDED RNA TO EXPOSE THE CODING
RP STRAND FOR TEMPLATING, INITIATES TRANSCRIPTION PREFERENTIALLY WITH
RP A PURINE.
CC -! CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -! SUBUNIT: Monomer.
CC -! SIMILARITY: Belongs to the phage and mitochondrial RNA polymerase
CC family.
CC -! DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/RNAPT7/".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01146; CAA24390.1; -;
CC EMBL; V01127; CAA24333.1; -;
CC EMBL; M38308; AAA32569.1; -;
CC PIR; A94615; RNBP17.
CC PDB; 4RNP; 03-DEC-97.
CC PDB; IARO; 21-OCT-98.
CC PDB; 1CEZ; 21-MAY-99.
CC PDB; IH38; 20-NOV-02.
CC PDB; IMSW; 15-NOV-02.
CC PDB; 1QLN; 04-FEB-00.
CC GO; GO:0019012; C:virion; NAS.
CC GO; GO:0003677; F:DNA binding; NAS.
CC GO; GO:0003899; F:DNA-directed RNA polymerase activity; NAS.
CC GO; GO:0006350; P:transcription; NAS.
CC InterPro; IPR002092; RNA_pol_phage.
CC Pfam; PF00940; RNA_pol; 1.
CC -----

DR PROSITE; PS00489; RNA_POL_PHASE_2; 1.
 DR PROSITE; PS00900; RNA_POL_PHASE_1; 1.
 KW Transferase; Transcription; DNA-directed RNA polymerase; 3D-structure.
 FT ACT_SITE 537 537
 FT ACT_SITE 631 631
 FT ACT_SITE 812 812
 FT MUTAGEN 172 172
 FT MUTAGEN 563 563
 FT MUTAGEN 571 571
 FT MUTAGEN 631 631
 FT MUTAGEN 631 631
 FT MUTAGEN 631 631
 FT MUTAGEN 636 636
 FT MUTAGEN 639 639
 FT MUTAGEN 646 646
 FT MUTAGEN 648 648
 FT CONFLICT 388 424
 FT
 FT CONFLICT 389 389
 FT CONFLICT 443 443
 FT CONFLICT 474 474
 FT CONFLICT 623 623
 FT CONFLICT 665 665
 FT TURN 13 14
 FT HELIX 15 27
 FT TURN 28 28
 FT HELIX 30 40
 FT HELIX 41 43
 FT TURN 44 44
 FT HELIX 48 58
 FT HELIX 77 92
 FT TURN 93 95
 FT TURN 101 102
 FT TURN 103 107
 FT TURN 108 108
 FT HELIX 111 125
 FT TURN 126 127
 FT STRAND 133 133
 FT HELIX 134 160
 FT TURN 161 161
 FT TURN 193 194
 FT HELIX 206 223
 FT STRAND 227 230
 FT STRAND 242 246
 FT HELIX 248 256
 FT HELIX 260 263
 FT STRAND 271 271
 FT STRAND 282 282
 FT STRAND 295 295
 FT STRAND 297 297
 FT HELIX 302 306
 FT TURN 307 307
 FT TURN 309 310
 FT STRAND 311 311
 FT HELIX 314 325
 FT STRAND 328 330
 FT HELIX 332 341
 FT TURN 342 342
 FT TURN 386 387
 FT HELIX 389 389
 FT TURN 408 408

Query Match 78.9%; Score 30; DB 1; Length 883;
 Best Local Similarity 60.0%; Pred. No. 3.4e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWH 5
 Db 201 WSSWH 205

RESULT 31
 FDHG_HAEIN

ID FDHG_HAEIN STANDARD; PRT; 1028 AA.
 AC P46449;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Formate dehydrogenase major subunit (EC 1.2.1.2) (Formate
 dehydrogenase alpha subunit) (FDH alpha subunit).
 GN FDHG OR HI0006
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-R4 / KW20 / ATCC 51907;
 RX MEDLINE=95350630; Pubmed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kiehl J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd";
 RL Science 269:496-512(1995).
 CC -!- FUNCTION: ALLOWS TO USE FORMATE AS MAJOR ELECTRON DONOR DURING
 CC ANAEROBIC RESPIRATION. SUBUNIT ALPHA POSSIBLY FORMS THE ACTIVE
 CC SITE.
 CC -!- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -!- COPACTOR: Molybdenum (molybdopterin) and selenocysteine. The
 CC active-site selenocysteine is encoded by the opal codon, UGA. May
 CC bind a 4Fe-4S cluster.
 CC -!- SUBUNIT: FORMATE DEHYDROGENASE IS A MEMBRANE-BOUND COMPLEX, FORMED
 CC BY SUBUNITS ALPHA, BETA AND GAMMA.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: ORTHOLOG OF BOTH E.COLI FDHG AND FDGO.
 CC -!- SIMILARITY: Belongs to the prokaryotic molybdopterin-containing
 CC oxidoreductase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; U32686; -; NOT_ANNOTATED_CDS.
 CC TIGR; HI0006; -;
 DR InterPro; IPR009010; Asp decarb fold.
 DR InterPro; IPR006443; Formate-dh-alph.
 DR InterPro; IPR006657; Mol dinuc bind.
 DR InterPro; IPR006963; Molybdop Fe4S4.
 DR InterPro; IPR006656; Molybdopterin.
 DR InterPro; IPR006655; Prok_Mboxred.
 DR InterPro; IPR006311; Tat.
 DR Pfam; PF04879; Molybdop_Fe4S4; 1.
 DR Pfam; PF00384; molybdopterin; 1.
 DR Pfam; PF01568; Molybdop binding; 1.
 DR TIGRFAMs; TIGR01553; formate-dh-alph; 1.
 DR TIGRFAMs; TIGR01409; TAT signal seq; 1.
 DR PROSITE; PS00551; MLYBDOPTERIN_PROK_1; 1.
 DR PROSITE; PS00490; MLYBDOPTERIN_PROK_2; FALSE_NEG.
 DR PROSITE; PS00932; MLYBDOPTERIN_PROK_3; 1.
 KW Oxidoreductase; Molybdenum; Selenocysteine; Selenium; NAD;
 KW Iron-sulfur; 4Fe-4S; Complete proteome.
 FT METAL 50 50 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 53 53 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 57 57 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 100 100 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

```
FT SE CYS 204 204 PROBABLE.
SQ SEQUENCE 1028 AA; 115403 MW; 295F185BB020EDF9 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 1028;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 639 WTQWHW 644

RESULT 32
POLG MCFA STANDARD; PRT; 3341 AA.
AC P33515;
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix
DE protein (Envelope protein M); Major envelope protein E; Nonstructural
DE proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase
DE (BC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
DE (NS5)].
OS Mosquito cell fusing agent (CFA flavivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OC NCBI_TaxID=31658;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92351550; PubMed=1322586;
RA Camisa-Parks H., Cisar L.A., Kane A., Stollar V.;
RT "The complete nucleotide sequence of cell fusing agent (CFA):
RT homology between the nonstructural proteins encoded by CFA and the
RT nonstructural proteins encoded by arthropod-borne flaviviruses.";
RL Virology 189:511-524(1992).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polyprotein, commonly with Asp or Glu in the P6
CC position, Cys or Thr in P1 and Ser or Ala in P1'.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -1- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
CC lipoprotein envelope. The envelope consists of two proteins:
CC protein M and glycoprotein E. The nucleocapsid is a complex of
CC protein C and mRNA.
CC -1- PTM: IN CFA THE PROCESSING OF PRM TO M EITHER DOES NOT OCCUR OR
CC OCCURS VERY INEFFICIENTLY.
CC -1- SIMILARITY: TO GENOME POLYPROTEIN OF OTHER FLAVIVIRUSES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M91671; AAA48509.1; -
DR PIR; A42996; A42996.
DR MEROPS; S07.UFW; -.
DR InterPro; IPR009003; Cys Ser tryptase.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000336; Flavi glycoprote.
DR InterPro; IPR001157; Flavi NS1.
DR InterPro; IPR000208; Flavi NS5.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR001850; peptidase S7.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR InterPro; IPR002877; RnmJ_PtsJ.
DR Pfam; PF00949; Flavi helicase; 1.
DR Pfam; PF00948; Flavi NS1; 1.
DR Pfam; PF00972; Flavi NS5; 1.

Pfam; PF01728; FtsJ; 1.
Pfam; PF00271; Helicase C; 1.
DR ProDom; PD001496; Flavi NS1; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Hydrolase; Helicase;
KW ATP-binding; Transmembrane; Nonstructural protein.
FT CHAIN 1 136
FT CHAIN 137 278
FT CHAIN 279 705
FT CHAIN 706 1095
FT CHAIN 1096 1327
FT CHAIN 1328 1451
FT CHAIN 1452 2028
FT CHAIN 2029 2196
FT CHAIN 2197 2454
FT CHAIN 2455 3341
FT DOMAIN 371 384
FT TRANSMEM 245 261
FT TRANSMEM 670 686
FT ACT_SITE 1506 1506
FT ACT_SITE 1530 1530
FT ACT_SITE 1589 1589
FT NP_BIND 1640 1647
FT SITE 1729 1732
FT CARBOHYD 157 157
FT CARBOHYD 243 243
FT CARBOHYD 339 339
FT CARBOHYD 399 399
FT CARBOHYD 411 411
FT CARBOHYD 575 575
FT CARBOHYD 611 611
FT CARBOHYD 794 794
FT CARBOHYD 896 896
FT CARBOHYD 993 993
FT CARBOHYD 1027 1027
SQ SEQUENCE 3341 AA; 373261 MW; BF0715F836F245ED CRC64;

Query Match 78.9%; Score 30; DB 1; Length 3341;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 1975 WLAWHW 1980

RESULT 33
VG73 BPML5 STANDARD; PRT; 85 AA.
AC Q05288;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Gene 73 protein (GP73).
DE 73.
GN Mycobacteriophage L5.
OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC L5-like viruses.
OC NCBI_TaxID=31757;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=93211282; PubMed=8459766;
RA Hatfull G.F., Sarkis G.J.;
RT "DNA sequence, structure and gene expression of mycobacteriophage L5;
RT a phage system for mycobacterial genetics.";
RL Mol. Microbiol. 7:395-405(1993).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z18946; CAA79449.1; -

DR PIR; S31018; S31018.

SQ SEQUENCE 85 AA; 10250 MW; 5032B5A1400FF4A7 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 85;
 Best Local Similarity 60.0%; Pred. No. 57;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWH 5

Db 42 WKTWH 46

RESULT 34

ID_VPX HV2D1 STANDARD; PRT; 112 AA.

AC P17760;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE VPX protein (X ORF protein) (Viral accessory protein).

GN VPX.

OS Human immunodeficiency virus type 2 (isolate D194) (HIV-2).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11713;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89184631; PubMed=2467304;

RA Kuehn H., von Briesen H., Dietrich U., Adamski M., Mix D.,

RA Biesert L., Kreutz R., Immanuel H., Henco K., Meichner C.,

RA Andreessen R., Gelderblom H., Ruebsamen-Waigmann H.;

RT "Molecular cloning of two west African human immunodeficiency virus

RT type 2 isolates that replicate well in macrophages: a Gambian

RT isolate, from a patient with neurologic acquired immunodeficiency

RT syndrome, and a highly divergent Ghanaian isolate.";

RL Proc. Natl. Acad. Sci. U.S.A. 86:2383-2387(1989).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=91045094; PubMed=2235509;

RA Kuehn H., Kreutz R., Ruebsamen-Waigmann H.;

RT "Nucleotide sequence of HIV-2D194, an isolate from a Gambian case of

RT 'neuro-AIDS', which showed excellent growth in macrophages.";

RL Nucleic Acids Res. 18:6142-6142(1990).

CC -!- MISCELLANEOUS: THIS ISOLATE IS FROM A GAMBIAN CASE OF

CC 'NEURO-AIDS'.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; J04542; AAA76843.1; -

DR EMBL; X52223; CAA36467.1; -

DR PIR; S12155; S12155.

DR HIV; J04542; VPX\$2D194.

DR InterPro: IPR000012; Retrov_Vpr/X.

DR Pfam: PF00522; VPR; 1.

DR PRINTS; PR00444; HIVVPRPX.

KW AIDS.

SQ SEQUENCE 112 AA; 12820 MW; 747AC8D908EA66B4 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 112;

Best Local Similarity 60.0%; Pred. No. 73;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWH 5

Db

53 WAYWH 57

RESULT 35

YP98 MYCTU

ID_YP98 MYCTU STANDARD; PRT; 164 AA.

AC Q50623;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein RV2598/MT2673.1/MB2629.

GN RV2598 OR MT2673.1 OR MTCY227.03C OR MB2629.

OS Mycobacterium tuberculosis, and

OS Mycobacterium bovis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773, 1765;

RN [1]

RP SEQUENCE FROM N.A.

RX SPECIES=M.tuberculosis; STRAIN=H37Rv;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence.";

RL Nature 393:537-544(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;

RX MEDLINE=22206494; PubMed=12218036;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,

RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;

RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and

RT laboratory strains.";

RL J. Bacteriol. 184:5479-5490(2002).

RN [3]

RP SEQUENCE FROM N.A.

RX SPECIES=M.bovis; STRAIN=AF2122/97;

RX MEDLINE=22709107; PubMed=12788972;

RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,

RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,

RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,

RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;

RT "The complete genome sequence of Mycobacterium bovis.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC

DR EMBL; Z77724; CAB01277.1; -

DR EMBL; AF007101; AAK46988.1; ALT INIT.

DR EMBL; BX248343; CAD94814.1; -

DR PIR; E70727; E70727.

DR TIGR; MT2673.1; -

DR Tuberculist; RV2598; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 164 AA; 17616 MW; 05F96ADB49103ED1 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 164;
 Best Local Similarity 60.0%; Pred. No. 1.e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
 DB 141 WRTH 145

RESULT 36

LSPA SERMA STANDARD; PRT; 170 AA.
 AC O2213;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lipoprotein signal peptidase (EC 3.4.23.36) (Prolipoprotein signal
 DE peptidase) (Signal peptidase II) (SPase II).
 GN LSPA.
 OS Serratia marcescens.
 OG Plasmid R471a.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Serratia.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Woodgate R.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: This protein specifically catalyzes the removal of
 CC signal peptides from prolipoproteins (By similarity).
 CC -1- CATALYTIC ACTIVITY: Cleavage of N-terminal leader sequences from
 CC membrane prolipoproteins. Hydrolyzes Xaa-Xbb-Xcc-|Cys, in which
 CC Xaa is hydrophobic (preferably Leu), Xbb is often Ser or Ala, Xcc
 CC is often Gly or Ala, and the Cys is alkylated on sulfur with a
 CC diacylglyceryl group.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: Belongs to peptidase family A8.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF027768; AAC82524.1; -;
 CC HAMAP; MF 00161; -;
 CC InterPro; IPR001872; Peptidase A8.
 CC Pfam; PF01252; Peptidase A8; 1;
 CC PRINTS; PR00781; LIPOSIGPTASE.
 CC PRODOM; PD004304; SigPase A8; 1.
 CC TIGRFAMs; TIGR00077; lspa; 1.
 CC PROSITE; PS00855; SPase II; 1.
 CC TRANSMEM; 1; POTENTIAL.
 CC TRANSMEM; 21; POTENTIAL.
 CC TRANSMEM; 43; POTENTIAL.
 CC TRANSMEM; 71; POTENTIAL.
 CC TRANSMEM; 131; POTENTIAL.
 CC ACT SITE 113 BY SIMILARITY.
 CC FT ACT SITE 140 BY SIMILARITY.
 CC FT ACT SITE 140 BY SIMILARITY.
 CC SEQUENCE 170 AA; 18815 MW; ECE1C7415B3C10FB CRC64;

Query Match 76.3%; Score 29; DB 1; Length 170;
 Best Local Similarity 50.0%; Pred. No. 1.e+02;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 6
 DB 128 WRDWH 133

RESULT 37

YDEJ_ECOLI STANDARD; PRT; 172 AA.
 AC P31131;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein ydeJ.
 GN YDEJ OR B1537.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93186717; PubMed=8383113;
 RA Cohen S.P., Haechler H., Levy S.B.;
 RT "Genetic and functional analysis of the multiple antibiotic
 RT resistance (mar) locus in Escherichia coli.";
 RL J. Bacteriol. 175:1484-1492(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori T.,
 RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
 RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubramam S.,
 RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
 RA Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 CC -1- SIMILARITY: BELONGS TO THE CINA FAMILY. STRONG, TO E.COLI YGAD.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M96235; -; NOT ANNOTATED_CDS.
 CC EMBL; AE000251; AAC74610.1; -;
 CC EMBL; D90796; BAA15227.1; -;
 CC EMBL; D90797; BAA15239.1; -;
 CC PIR; D64908; D64908.
 CC EcoGene; EGI1645; ydeJ.
 CC HAMAP; MF 00226; atypical; 1.
 CC InterPro; IPR008136; CINA_C.
 CC Pfam; PF02464; CINA; 1.
 CC TIGRFAMs; TIGR00199; cina cterm; 1.
 CC Hypothetical protein; Complete proteome.
 CC SEQUENCE 172 AA; 18321 MW; 700A7A95C967C868 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 172;
 Best Local Similarity 60.0%; Pred. No. 1.e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
 DB 132 WFAWH 136

```

RESULT 38
YC07 METJA
ID YC07 METJA STANDARD; PRT; 226 AA.
AC Q89604;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical acetyltransferase MJ1207 (EC 2.3.1.-).
GN MJ1207.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
SEQUENCE FROM N.A.
STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
MEDLINE=96337999; PubMed=868087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervilange A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RA jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE ACETYLTRANSFERASE FAMILY. STRONG, TO
CC A.FULGIDUS AF0521.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67562; AAB99211.1; -.
CC PIR; F64450; F64450.
CC TIGR; MJ1207; -.
CC InterPro; IPR000182; GCSN5acetyl_trans.
CC Pfam; PF00583; Acetyltransf; 1.
CC Hypothetical protein; Transferase; Acyltransferase; Complete proteome.
CC
CC QUERY MATCH 76.3%; Score 29; DB 1; Length 226;
CC Best Local Similarity 50.0%; Pred. No. 1.4e+02;
CC Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 1 WXXWHF 6
CC | | |
CC 109 WARWYF 114
CC
CC DB
CC
CC RESULT 39
LIPB THETN
ID LIPB THETN STANDARD; PRT; 228 AA.
AC Q8R9E0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lipote-protein ligase B (EC 6.-.-.-) (Lipoate biosynthesis protein
DE B).
DE B.
GN LIPB OR TTE1673.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
SEQUENCE FROM N.A.
STRAIN=MB4 / JCM 11007;
MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- FUNCTION: Involved in the attachment of lipoyl groups to proteins,
CC by creating an amide linkage that joins the free carboxyl group of
CC lipoic acid to the epsilon-amino group of a specific lysine
CC residue in lipoylated proteins (By similarity).
CC -!- PATHWAY: Lipoate biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the lipB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE013121; AAM24874.1; -.
CC HAMAP; MF 00013; -.
CC InterPro; IPR004143; BPL_LipA_LipB.
CC InterPro; IPR000544; Lipoate_B.
CC Pfam; PF03099; BPL_LipA_LipB; 1.
CC ProDom; PD006086; Lipoate_B; 1.
CC TIGRFAMs; TIGR00214; lipB; 1.
CC PROSITE; PS01313; LIPB; 1.
CC Ligase; Complete proteome.
CC
CC QUERY MATCH 76.3%; Score 29; DB 1; Length 228;
CC Best Local Similarity 60.0%; Pred. No. 1.4e+02;
CC Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 1 WXXWH 5
CC | | |
CC 151 WITWH 155
CC
CC DB
CC
CC RESULT 40
HIS4 SYN7
ID HIS4 SYN7 STANDARD; PRT; 254 AA.
AC Q8GJM0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]
DE imidazole-4-carboxamide isomerase (EC 5.3.1.16)
DE DE (Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide
DE isomerase).
DE HIS4 OR SEN0020.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
SEQUENCE FROM N.A.
RA Holtman C.K., Sandoval P., Chen Y., Socias T., McMurtry S.,
RA Gonzalez A., Salinas I., Golden S.S., Youderian P.;
RA Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]
CC imidazole-4-carboxamide + 5-phosphoribosyl-1-ylamino]methylideneamino] -> 5-
CC [(5-phospho-1-deoxyribulos-1-ylamino)methylideneamino]-1-(5-
CC phosphoribosyl)imidazole-4-carboxamide.
CC -!- PATHWAY: Histidine biosynthesis; fourth step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the hisA / hisF family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AY157498; ANA46174.1; -.
DR HAMAP; MF_01014; -; 1.
DR InterPro; IPR003009; FMN enzyme.
DR InterPro; IPR006063; His4.
DR InterPro; IPR006062; His biosynth.
DR Pfam; PF00977; His biosynth; 1.
DR TIGRFAMs; TIGR00007; TIGR00007; 1.
KW Isomerase; Histidine biosynthesis.
SQ SEQUENCE 254 AA; 26820 MW; 16D1A159A2C45514 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 254;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
Db 167 WARWH 171

RESULT 41

S1C2_HUMAN STANDARD; PRT; 302 AA.
ID - S1C2_HUMAN STANDARD; PRT; 302 AA.
AC O75897;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfotransferase 1C2 (EC 2.8.2.-) (SULT1C) (SULT1C#2).
GN SULT1C2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=99069375; PubMed=9852044;
RA Sakakibara Y., Yanagisawa K., Katafuchi J., Ringer D.P., Takami Y.,
RA Nakayama T., Suiko M., Liu M.-C.;
RT "Molecular cloning, expression, and characterization of novel human
RT SULTC sulfotransferases that catalyze the sulfonation of
RT N-hydroxy-2-acetylaminofluorene."
RL J. Biol. Chem. 273:33929-33935(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20247255; PubMed=10783263;
RA Freimuth R.R., Raftogianis R.B., Wood T.C., Moon E., Kim U.-J., Xu J.,
RA Sciallano M.J., Weinshilboum R.M.;
RT "Human sulfotransferases SULT1C1 and SULT1C2: cDNA characterization,
RT gene cloning, and chromosomal localization."
RL Genomics 65:157-165(2000).
CC -1- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters. May be
CC involved in the activation of carcinogenic hydroxylamines. Shows
CC activity towards p-nitrophenol and N-hydroxy-2-acetylami-
CC fluorene (N-OH-2AFL).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: Expressed at high levels in fetal lung and
CC kidney and at low levels in fetal heart, adult kidney, ovary and
CC spinal chord.
CC -1- SIMILARITY: Belongs to the sulfotransferase family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AF055584; AAC95519.1; -.
DR EMBL; AF186263; AAF73810.1; -.
DR HSSP; P50224; 1CJM.
DR Genew; HGNC:11457; SULT1C2.
DR GO; GO:0008146; F:sulfotransferase activity; NAS.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
FT CONFLICT 5 5 E -> D (IN REF. 2).
SQ SEQUENCE 302 AA; 35534 MW; DD2F4865B2E3B56D CRC64;

Query Match 76.3%; Score 29; DB 1; Length 302;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
Db 176 WGSWH 180

RESULT 42

FG11_HUMAN STANDARD; PRT; 312 AA.
ID - FG11_HUMAN STANDARD; PRT; 312 AA.
AC Q08830; Q96KW6; Q96QW6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen-like protein 1 precursor (Hepatocyte-derived fibrinogen-
DE related protein 1) (HFRP-1) (Hepasocin) (HP-041).
GN FG11 OR HFRP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93290661; PubMed=8390249;
RA Yamamoto T., Gotoh M., Sasaki H., Terada M., Kitajima M.,
RA Hirohashi S.;
RT "Molecular cloning and initial characterization of a novel
RT fibrinogen-related gene, HFRP-1."
RL Biochem. Biophys. Res. Commun. 193:681-687(1993).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=21363035; PubMed=11470158;
RA Hara H., Yoshimura H., Uchida S., Toyoda Y., Aoki M., Sakai Y.,
RA Morimoto S., Shiohara K.;
RT "Molecular cloning and functional expression analysis of a cDNA for
RT human hepasocin, a liver-specific protein with hepatocyte mitogenic
RT activity."
RL Biochim. Biophys. Acta 1520:45-53(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22398257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeng B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Groomwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.L., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Has hepatocyte mitogenic activity.
CC -!- SUBUNIT: Homodimer (Probable).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Liver specific.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D14446; BA003336.1; -;
DR EMBL; D87342; BAB70690.1; -;
DR EMBL; BC007047; AAH07047.1; -;
DR PIR; JN0596; JN0596.
DR HSSP; P02671; 1FZD.
DR Genbank; HGNC:3695; FGL1.
DR MIM; 605776; -;
DR GO; GO:0005577; C:fibrinogen complex; TAS.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW SIGNAL.
FT SIGNAL 1 22
FT CHAIN 23 312 FIBRINOGEN-LIKE PROTEIN 1.
FT DOMAIN 78 305 FIBRINOGEN C-TERMINAL.
FT DISULFID 26 26 INTERCHAIN (POTENTIAL).
FT DISULFID 83 112 BY SIMILARITY.
FT DISULFID 248 261 BY SIMILARITY.
FT CONFLICT 15 15 I -> T (IN REF. 3).
FT CONFLICT 69 69 N -> D (IN REF. 1).
FT CONFLICT 72 72 I -> V (IN REF. 1).
FT CONFLICT 105 105 P -> L (IN REF. 2).
SQ SEQUENCE 312 AA; 36391 MW; 26BC82124E6660C2 CRC64;
Query Match 76.3%; Score 29; DB 1; Length 312;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 WXXWH 5
Db 285 WYTW 289
RESULT 43
TYSY LACCA
AC P00469; STANDARD; PRT; 316 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thymidylate synthase (EC 2.1.1.45) (TS) (TSase).
GN THVA.
OS Lactobacillus casei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1582;
FN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88283342; PubMed=2840247;
RA Pinter K., Davidsson V.J., Santi D.V.;
"Cloning, sequencing, and expression of the Lactobacillus casei
thymidylate synthase gene.";
DNA 7:235-241(1988).
[2]
RP SEQUENCE.
RX MEDLINE=79109713; PubMed=105005;
RA Maley G.F., Bellisario R.L., Guarino D.U., Maley F.;
RT "The primary structure of Lactobacillus casei thymidylate synthetase.
RT III. The use of 2-(2-nitrophenylsulfonyl)-3-methyl-3-bromindoline
RT and limited tryptic peptides to establish the complete amino acid
RT sequence of the enzyme.";
RL J. Biol. Chem. 254:1301-1304(1979).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS).
RX MEDLINE=91219394; PubMed=2128651;
RA Perry K.M., Fauman E.B., Finer-Moore J.S., Montfort W.R., Maley G.F.,
RA Maley F., Stroud R.M.;
RT "Plastic adaptation toward mutations in proteins: structural
RT comparison of thymidylate synthases.";
RL Proteins 8:315-333(1990).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.55 ANGSTROMS).
RX MEDLINE=93381690; PubMed=8371269;
RA Finer-Moore J.S., Fauman E.B., Foster P.G., Perry K.M., Santi D.V.,
RA Stroud R.M.;
RT "Refined structures of substrate-bound and phosphate-bound
RT thymidylate synthase from Lactobacillus casei.";
RL J. Mol. Biol. 232:1101-1116(1993).
RN [5]
RP MUTAGENESIS OF TYR-82.
RX MEDLINE=98044713; PubMed=9383465;
RA Kealey J.T., Eckstein J., Santi D.V.;
RT "Role of the conserved tryptophan 82 of Lactobacillus casei
RT thymidylate synthase.";
RL Chem. Biol. 2:609-614(1995).
CC -!- FUNCTION: Provides the sole de novo source of dTMP for DNA
CC biosynthesis.
CC -!- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP =
CC dihydrofolate + dTMP.
CC -!- PATHWAY: Deoxyribonucleotide biosynthesis.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the thymidylate synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M19653; AAA25255.1; ALT_INIT.
DR PIR; A29817; SYLBT.
DR PDB; 1TMS; 15-JAN-92.
DR PDB; 1TDA; 15-JUL-93.
DR PDB; 1TDB; 15-JUL-93.
DR PDB; 1TDC; 15-JUL-93.
DR PDB; 1TDD; 15-JUL-93.
DR PDB; 1TDM; 20-AUG-97.
DR PDB; 1THY; 31-OCT-93.
DR PDB; 1TSV; 08-MAR-96.
DR PDB; 1TSM; 08-MAR-96.
DR PDB; 1TSX; 08-MAR-96.
DR PDB; 1TSY; 08-MAR-96.
DR PDB; 1TSZ; 08-MAR-96.
DR PDB; 1LCA; 15-OCT-95.
DR PDB; 1LCB; 15-OCT-95.
DR PDB; 1LCE; 15-OCT-95.
DR PDB; 1NJA; 11-JUL-96.
DR PDB; 1NJB; 14-OCT-96.
DR PDB; 1NJC; 11-JUL-96.
DR PDB; 1NJD; 11-JUL-96.

```

DR PDB; 1NJE; 11-JUL-96.
DR PDB; 1VZA; 12-MAR-97.
DR PDB; 1VZB; 12-MAR-97.
DR PDB; 1VZC; 12-MAR-97.
DR PDB; 1VZD; 12-MAR-97.
DR PDB; 1VZE; 12-MAR-97.
DR PDB; 1VMF; 28-JAN-98.
DR PDB; 1VMG; 28-JAN-98.
DR PDB; 1VMH; 28-JAN-98.
DR PDB; 1VMI; 28-JAN-98.
DR PDB; 1TSL; 17-JUN-98.
DR PDB; 1TSM; 17-JUN-98.
DR PDB; 1TSM; 17-JUN-98.
DR PDB; 1TVU; 28-JAN-98.
DR PDB; 1TVV; 28-JAN-98.
DR PDB; 1TVW; 28-JAN-98.
DR PDB; 1BO7; 19-AUG-98.
DR PDB; 1BO8; 19-AUG-98.
DR PDB; 1BP0; 19-AUG-98.
DR PDB; 1BP6; 19-AUG-98.
DR PDB; 1BPJ; 19-AUG-98.
DR HAMAP; MF_00008; -.
DR InterPro; IPR000398; Thymidylat synth.
DR Pfam; PF00303; thymidylat synt; 1.
DR PRINTS; P00108; THYMSDNTASE.
DR PRODom; P001180; Thymidylat synt; 2.
DR PROSITE; PS00091; THYMIDYLATE SYNTHASE; 1.
KW Transferase; Methyltransferase; Nucleotide biosynthesis; 3D-structure.
FT ACT SITE 198 199
FT TURN 2 3
FT HELIX 4 16
FT STRAND 18 20
FT TURN 23 24
FT STRAND 28 39
FT TURN 40 41
FT HELIX 54 65
FT TURN 66 67
FT STRAND 70 70
FT TURN 71 71
FT HELIX 72 75
FT TURN 76 78
FT TURN 81 82
FT HELIX 83 91
FT TURN 92 92
FT TURN 94 95
FT TURN 102 103
FT HELIX 104 107
FT HELIX 108 110
FT TURN 112 113
FT HELIX 114 131
FT HELIX 133 139
FT TURN 140 140
FT STRAND 141 141
FT HELIX 146 151
FT STRAND 153 154
FT TURN 156 157
FT STRAND 160 161
FT HELIX 163 173
FT TURN 175 176
FT TURN 178 179
FT STRAND 181 183
FT TURN 187 189
FT HELIX 190 192
FT STRAND 199 207
FT TURN 208 209
FT STRAND 210 221
FT TURN 222 225
FT HELIX 226 244
FT TURN 245 245
FT STRAND 247 261
FT HELIX 262 264
FT HELIX 265 271
FT TURN 272 273
FT STRAND 281 284

Query Match 76.3%; Score 29; DB 1; Length 316;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 150 WRAWH 154

RESULT 44
LYCL_YARLI
ID LYCL_YARLI STANDARD; PRT; 391 AA.
AC P41929;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lysine acetyltransferase (EC 2.3.1.32) (Lysine N(6)-acetyltransferase)
DE (LAT).
GN LYCL.
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 20460 / W29;
RX MEDLINE=94363731; PubMed=8082161;
RA Beckerich J.-M., Lambert M., Gaillardin C.;
RT "LYCL is the structural gene for lysine N-6-acetyl transferase in
RL yeast.";
RL Curr. Genet. 25:24-29(1994).
CC -|- CATALYTIC ACTIVITY: Acetyl phosphate + L-lysine = phosphate +
CC N(6)-acetyl-L-lysine.
CC -|- PATHWAY: Lysine metabolism; first step.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X63548; CAA45112.1; -
DR PIR; S39816; S39816.
KW Transferase; Acyltransferase.
SQ SEQUENCE 391 AA; 43306 MW; DAC927FC5FD4119F CRC64;

Query Match 76.3%; Score 29; DB 1; Length 391;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 278 WRAWH 282

RESULT 45
ALGI_PSEPK
ID ALGI_PSEPK STANDARD; PRT; 485 AA.
AC Q88ND2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable poly(beta-D-mannuronate) O-acetylase (EC 2.3.1.-) (Alginate
DE biosynthesis protein algi).
GN ALGI OR PP1280.
OS Pseudomonas putida (strain KT2440).

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]_TaxID=294;
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weiner C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holzapple E., Scanlan D., Tran K., Moazzez A.,
RA Uckerback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Helm S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Dueterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
CC -!- FUNCTION: Together with algJ and algF, forms an inner membrane
CC complex which probably interacts with the alginate polymerization-
CC transport complex and adds acetyl groups at the O-2 and O-3
CC positions of polymannuronic acid. Acetylation of alginate is
CC important for the architecture of biofilms and increases the
CC ability of alginate to act as a defense barrier (By similarity).
CC -!- PATHWAY: Alginate biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: Belongs to the membrane-bound acyltransferase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF016778; AAN66904.1; -.
DR TIGR; PP1280; -.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
KW Alginate biosynthesis; Transferase; Acyltransferase; Inner membrane;
KW Transmembrane; Complete proteome.
FT TRANSMEM 7 24 POTENTIAL.
FT TRANSMEM 39 61 POTENTIAL.
FT TRANSMEM 78 100 POTENTIAL.
FT TRANSMEM 115 137 POTENTIAL.
FT TRANSMEM 150 172 POTENTIAL.
FT TRANSMEM 312 334 POTENTIAL.
FT TRANSMEM 360 382 POTENTIAL.
FT TRANSMEM 402 424 POTENTIAL.
FT TRANSMEM 461 483 POTENTIAL.
FT ACT_SITE 322 322 POTENTIAL.
SQ SEQUENCE 485 AA; 55451 MW; D3E30C528EB92A1 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 485;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 331 WGAWH 335

RESULT 46
ALGI_PSEFL
ID ALGI_PSEFL STANDARD; PRT; 495 AA.
AC PS9789;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable poly(beta-D-mannuronate) O-acetylase (EC 2.3.1.1-) (Alginate
DE biosynthesis protein algI).

GN ALGI.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]_TaxID=294;
RP SEQUENCE FROM N.A.
RX STRAIN=NCIMB 10525;
RX PubMed=12775688;
RA Gimnesstad M., Sletta H., Ertesvaag H., Bakkevig K., Jain S.,
RA Suh S.-J., Skjaak-Braek G., Ellingsen T.E., Ohman D.E., Valla S.;
RT "The Pseudomonas fluorescens AlgG protein, but not its mannuronan C-5-
RT epimerase activity, is needed for alginate polymer formation."
RL J. Bacteriol. 185:3515-3523(2003).
CC -!- FUNCTION: Together with algJ and algF, forms an inner membrane
CC complex which probably interacts with the alginate polymerization-
CC transport complex and adds acetyl groups at the O-2 and O-3
CC positions of polymannuronic acid. Acetylation of alginate is
CC important for the architecture of biofilms and increases the
CC ability of alginate to act as a defense barrier (By similarity).
CC -!- PATHWAY: Alginate biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: Belongs to the membrane-bound acyltransferase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF527790; AAP46697.1; -.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
KW Alginate biosynthesis; Transferase; Acyltransferase; Inner membrane;
KW Transmembrane.
FT TRANSMEM 7 24 POTENTIAL.
FT TRANSMEM 39 61 POTENTIAL.
FT TRANSMEM 101 123 POTENTIAL.
FT TRANSMEM 136 158 POTENTIAL.
FT TRANSMEM 296 318 POTENTIAL.
FT TRANSMEM 344 366 POTENTIAL.
FT TRANSMEM 387 409 POTENTIAL.
FT TRANSMEM 471 493 POTENTIAL.
FT ACT_SITE 307 307 POTENTIAL.
SQ SEQUENCE 495 AA; 56185 MW; AB14A15B9757DED CRC64;

Query Match 76.3%; Score 29; DB 1; Length 495;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
DB 316 WGAWH 320

RESULT 47
ALGI_AZOVI
ID ALGI_AZOVI STANDARD; PRT; 499 AA.
AC O52156;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable poly(beta-D-mannuronate) O-acetylase (EC 2.3.1.1-) (Alginate
DE biosynthesis protein algI).
GN ALGI.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]_TaxID=354;

```

```

SEQUENCE FROM N.A.
RP STRAIN=ATCC 9046;
RX MEDLINE=92282527; PubMed=10352233;
RA Vazquez A., Moreno S., Guzman J., Alvarado A., Espin G.;
RT "Transcriptional organization of the Azotobacter vinelandii algXLVIFA
RL Gene 232:217-222(1999).
CC
CC -!- FUNCTION: Together with algV and algF, forms an inner membrane
CC complex which probably interacts with the alginate polymerization-
CC transport complex and adds acetyl groups at the O-2 and O-3
CC positions of polymannuronic acid. Acetylation of alginate
CC increases cyst resistance to desiccation.
CC -!- PATHWAY: Algininate biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: Belongs to the membrane-bound acyltransferase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF027499; AAC04568.1; -
CC InterPro; IPR004299; MBOAT_fam.
CC Pfam; PF03062; MBOAT; 1.
CC Algininate biosynthesis; Transferase; Acyltransferase; Inner membrane;
CC Transmembrane.
CC TRANSMEM 7 25 POTENTIAL.
CC TRANSMEM 40 62 POTENTIAL.
CC TRANSMEM 78 100 POTENTIAL.
CC TRANSMEM 115 137 POTENTIAL.
CC TRANSMEM 150 172 POTENTIAL.
CC TRANSMEM 239 261 POTENTIAL.
CC TRANSMEM 312 334 POTENTIAL.
CC TRANSMEM 354 373 POTENTIAL.
CC TRANSMEM 380 397 POTENTIAL.
CC TRANSMEM 407 429 POTENTIAL.
CC TRANSMEM 475 497 POTENTIAL.
CC ACT SITE 322 322 POTENTIAL.
CC SEQUENCE 499 AA; 56386 MW; CB874AE4DF08F63C CRC64;

Query Match 76.3%; Score 29; DB 1; Length 499;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
Db 331 WGAWH 335

RESULT 48
ALGI_PSESM
ID ALGI_PSESM STANDARD; PRT; 518 AA.
AC Q88706;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable poly(beta-D-mannuronate) O-acetylase (EC 2.3.1.-) (Algininate
DE biosynthesis protein algI).
GN ALGI OR PSPT01235.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,

```

```

RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,
RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC
CC -!- FUNCTION: Together with algV and algF, forms an inner membrane
CC complex which probably interacts with the alginate polymerization-
CC transport complex and adds acetyl groups at the O-2 and O-3
CC positions of polymannuronic acid. Acetylation of alginate is
CC important for the architecture of biofilms and increases the
CC ability of alginate to act as a defense barrier (By similarity).
CC -!- PATHWAY: Algininate biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: Belongs to the membrane-bound acyltransferase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE016860; AA054760.1; -
CC TIGR; PSPT0235; -
CC InterPro; IPR004299; MBOAT_fam.
CC Pfam; PF03062; MBOAT; 1.
CC Algininate biosynthesis; Transferase; Acyltransferase; Inner membrane;
CC Transmembrane; Complete proteome.
CC TRANSMEM 2 24 POTENTIAL.
CC TRANSMEM 39 61 POTENTIAL.
CC TRANSMEM 78 100 POTENTIAL.
CC TRANSMEM 115 137 POTENTIAL.
CC TRANSMEM 150 172 POTENTIAL.
CC TRANSMEM 319 341 POTENTIAL.
CC TRANSMEM 354 373 POTENTIAL.
CC TRANSMEM 402 424 POTENTIAL.
CC TRANSMEM 493 515 POTENTIAL.
CC ACT SITE 322 322 POTENTIAL.
CC SEQUENCE 518 AA; 58606 MW; 4621789C6FACD9B5 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 518;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
Db 331 WGAWH 335

RESULT 49
ALGI_PSEAE
ID ALGI_PSEAE STANDARD; PRT; 520 AA.
AC Q51392; Q9HY67;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable poly(beta-D-mannuronate) O-acetylase (EC 2.3.1.-) (Algininate
DE biosynthesis protein algI).
GN ALGI OR PA3548.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RC SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RP STRAIN=FRD1,

```

RX MEDLINE=96218684; PubMed=8636017;
RA Franklin M.J., Ohman D.E.;
RT "Identification of algI and algJ in the Pseudomonas aeruginosa
RT alginate biosynthetic gene cluster which are required for alginate O
acetylation.";
RL J. Bacteriol. 178:2186-2195(1996).
RN (2)
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PAOI;
RA MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
RN (3)
RP SUBCELLULAR LOCATION.
RX STRAIN=FRD1;
RA MEDLINE=21999262; PubMed=12003941;
RA Franklin M.J., Ohman D.E.;
RT "Mutant analysis and cellular localization of the AlgI, AlgJ, and
RT AlgF proteins required for O acetylation of alginate in Pseudomonas
aeruginosa.";
RL J. Bacteriol. 184:3000-3007(2002).
RN (4)
RP ROLE IN BIOFILM FORMATION.
RX STRAIN=FRD1;
RA MEDLINE=21142532; PubMed=11208804;
RA Nivens D.E., Ohman D.E., Williams J., Franklin M.J.;
RT "Role of alginate and its O acetylation in formation of Pseudomonas
aeruginosa microcolonies and biofilms.";
RL J. Bacteriol. 183:1047-1057(2001).
RN (5)
RP ROLE IN RESISTANCE TO PHAGOCYTOSIS.
RX STRAIN=FRD1;
RA MEDLINE=21117014; PubMed=11179370;
RA Pier G.B., Coleman F., Groot M., Franklin M., Ohman D.E.;
RT "Role of alginate O acetylation in resistance of mucoid Pseudomonas
aeruginosa to opsonic phagocytosis.";
RL Infect. Immun. 69:1895-1901(2001).
CC -|- FUNCTION: Together with algJ and algF, forms an inner membrane
complex which probably interacts with the alginate polymerization-
transport complex and adds acetyl groups at the O-2 and O-3
positions of polymannuronic acid. Acetylation of alginate is
important for the architecture of biofilms and increases
resistance to opsonic killing in the host.
CC -|- PATHWAY: Alginate biosynthesis.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -|- SIMILARITY: Belongs to the membrane-bound acyltransferase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U50202; AAB09781.1; --
CC ENBL; AE004775; AAG06936.1; --
CC PIR; A83203; A83203.
CC InterPro; IPR004299; MBOAT_fam.
CC Pfam; PF03062; MBOAT; 1.
CC Alginate biosynthesis; Transferase; Acyltransferase; Inner membrane;
Transmembrane; Complete proteome.
TRANSMEM 7 24
TRANSMEM 39 61
TRANSMEM 78 100
TRANSMEM 115 137
FT

FT TRANSMEM 150 172
FT TRANSMEM 239 261
FT TRANSMEM 311 333
FT TRANSMEM 353 375
FT TRANSMEM 402 424
FT TRANSMEM 483 505
FT ACT SITE 322 322
FT CONFLICT 360 361
FT CONFLICT 403 403
SQ SEQUENCE 520 AA; 58746 MW; 29146A26A2C6738B CRC64;
Query Match 76.3%; Score 29; DB 1; Length 520;
Best Local Similarity 60.0%; Pred.No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 WXXWH 5
Db 331 WGAWH 335
RESULT 50
PM21 LYCES STANDARD; PRT; 550 AA.
AC P05607; Q43144; Q43777;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pectinesterase 2 precursor (EC 3.1.1.11) (Pectin methylsterase 2)
DE (PE 2).
DE PME2.1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamnids; Solanales; Solanaceae; Solanum.
OC NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ailsa Craig; TISSUE=Pericarp;
RX MEDLINE=94289653; PubMed=8018878;
RA Hall L.N., Bird C.R., Picton S., Tucker G.A., Seymour G.B.,
RA Grierson D.;
RT "Molecular characterisation of cDNA clones representing pectin
esterase isozymes from tomato.";
RL Plant Mol. Biol. 45:313-318(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Rutgers; TISSUE=Fruit;
RA Turner L.A., Kausch K.D., Handa A.K.;
RT "Isolation and nucleotide sequence of two cDNAs corresponding to
tomato fruit pectin methylsterase genes.";
RL (In) Plant Gene Register FGR96-035.
RN [3]
RP SEQUENCE OF 177-550 FROM N.A.
RX MEDLINE=88225091; PubMed=3371355;
RA Ray J., Knapp J., Grierson D., Bird C., Schuch W.;
RT "Identification and sequence determination of a cDNA clone for tomato
pectin esterase.";
RL Eur. J. Biochem. 174:119-124(1988).
CC -|- FUNCTION: Pectinesterase may play a role in cell wall metabolism
during fruit growth and development prior to ripening and may be
required for preparing cell walls for softening by
polygalacturonase during fruit ripening.
CC -|- CATALYTIC ACTIVITY: Pectin + N H(2)O = N methanol + pectate.
CC -|- SUBCELLULAR LOCATION: Cell wall.
CC -|- DEVELOPMENTAL STAGE: In ripening fruit.
CC -|- SIMILARITY: Belongs to the pectinesterase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC


```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74639; CAA52704.1; -.
DR EMBL; U50985; AAB67739.1; -.
DR EMBL; X07910; CAA30746.1; -.
DR EMBL; A15983; CAA01257.1; -.
DR FIR; S46528; S46528.
DR InterPro; IPR000070; Pectinesterase.
DR InterPro; IPR006501; PME_inhib.
DR InterPro; IPR007186; PME_I.
DR Pfam; PF01095; Pectinesterase; 1.
DR Pfam; PF04043; PME1; 1.
DR TIGREAFMS; TIGR01614; PME_inhib; 1.
DR PROSITE; PS00800; PECTINESTERASE 1; 1.
DR PROSITE; PS00503; PECTINESTERASE 2; 1.
KW Hydrolase; Aspartyl esterase; Cell wall; Fruit ripening; Zymogen;
KW Signal; Multigene family.
FT SIGNAL 1 ? POTENTIAL.
FT PROPEP 233 ? POTENTIAL.
FT CHAIN 234 550 PECTINESTERASE 2.
FT DISULFID 331 358 BY SIMILARITY.
FT DISULFID 399 433 BY SIMILARITY.
FT ACT_SITE 365 365 BY SIMILARITY.
FT ACT_SITE 386 386 BY SIMILARITY.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 377 378 RD -> QS (IN REF. 3).
FT CONFLICT 468 468 Y -> S (IN REF. 3).
FT CONFLICT 515 515 V -> C (IN REF. 2).
FT CONFLICT 550 550 D -> DYSDIKLLFVYVTRHL (IN REF. 3).
SQ SEQUENCE 550 AA; 60500 MW; 7AB87CC37E849DE7 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 550;
Best Local Similarity 60.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWH 5
Db 478 WAEWH 482

```

Search completed: June 10, 2004, 10:48:42
Job time : 11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2004, 10:43:04 ; Search time 33 Seconds
(without alignments)
57.367 Million cell updates/sec

Title: US-09-912-414-11
Perfect score: 38
Sequence: 1 WXXWHF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database :

SPTREMBL 25.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	94.7	302	2 Q8GJ78	Q8gJ78 mycobacteri
2	35	92.1	241	17 Q59123	Q59123 pyrococcus
3	35	92.1	256	10 Q9FIN2	Q9fin2 arabidopsis
4	35	92.1	258	16 Q8NLA5	Q8nla5 corynebacte
5	35	92.1	276	17 Q8PU18	Q8pu18 methanosarc
6	35	92.1	289	17 Q8TRL3	Q8trl3 methanosarc
7	35	92.1	314	16 Q9L120	Q9l120 streptomyce
8	35	92.1	325	10 Q8LDW2	Q8ldw2 arabidopsis
9	35	92.1	325	10 Q8G242	Q8g242 bruceella su
10	35	92.1	377	16 Q8G242	Q8g242 arabidopsis
11	35	92.1	381	2 Q93TQ4	Q93tq4 agrobacteri
12	35	92.1	386	16 Q8UH89	Q8uh89 agrobacteri
13	35	92.1	394	10 Q9AW80	Q9aw80 guillardia
14	35	92.1	403	16 Q8YFS6	Q8yfs6 bruceella me
15	35	92.1	419	16 Q7UNAS	Q7una5 rhodopirell
16	35	92.1	421	10 Q9SBU4	Q9sbu4 chlamydomon

17	35	92.1	454	10 Q9ZRW0	Q9zrw0 cicer ariet
18	35	92.1	469	16 Q7UUM9	Q7ujm9 rhodopirell
19	35	92.1	497	3 Q94514	Q94514 schizosacch
20	35	92.1	499	10 Q9LTV0	Q9ltv0 arabidopsis
21	35	92.1	510	16 Q8XQC5	Q8xqc5 ralsstonia s
22	35	92.1	522	10 Q9SGT7	Q9sgt7 arabidopsis
23	35	92.1	525	16 Q8XQC8	Q8xqc8 ralsstonia s
24	35	92.1	549	10 Q8H2U5	Q8h2u5 oryza sativ
25	35	92.1	594	5 Q8II13	Q8i1i3 plasmodium
26	34	89.5	187	16 Q9CLS4	Q9cls4 pasteurella
27	34	89.5	190	4 Q96N05	Q96n05 homo sapien
28	34	89.5	201	10 Q9C600	Q9c600 arabidopsis
29	34	89.5	202	10 Q9AXY3	Q9axy3 brassica ca
30	34	89.5	213	10 Q9AXY5	Q9axy5 brassica na
31	34	89.5	216	10 Q9AXY2	Q9axy2 brassica na
32	34	89.5	222	10 Q9AXY4	Q9axy4 brassica ol
33	34	89.5	230	16 Q53604	Q53604 mycobacteri
34	34	89.5	230	16 Q7U2Y3	Q7u2y3 mycobacteri
35	34	89.5	297	16 Q8EWI8	Q8ewi8 mycoplasma
36	34	89.5	310	2 Q9EXF1	Q9exf1 listeria mo
37	34	89.5	415	16 Q8RD64	Q8rd64 thermoanaer
38	34	89.5	423	10 Q8H0T7	Q8h0t7 arabidopsis
39	34	89.5	423	16 Q927H4	Q927h4 listeria in
40	34	89.5	423	16 Q8Y412	Q8y412 listeria mo
41	34	89.5	424	10 Q48663	Q48663 chlamydomon
42	34	89.5	452	16 Q8ZL19	Q8z1l9 salmonella
43	34	89.5	452	16 Q8Z2K1	Q8z2k1 salmonella
44	34	89.5	462	16 Q8X5T1	Q8x5t1 escherichia
45	34	89.5	462	16 Q8FCM4	Q8fcm4 escherichia
46	34	89.5	465	16 Q8A0B4	Q8a0b4 bacteroides
47	34	89.5	469	5 Q8IJV7	Q8ijv7 plasmodium
48	34	89.5	472	16 Q8YMN8	Q8ymn8 anabaena sp
49	34	89.5	473	5 Q9BRU7	Q9brm7 leishmania
50	34	89.5	473	11 Q70396	Q70396 mus musculu
51	34	89.5	475	17 Q9HSR8	Q9hsr8 halobacteri
52	34	89.5	479	16 Q88OK9	Q88ok9 pseudomonas
53	34	89.5	485	10 Q80401	Q80401 oryza sativ
54	34	89.5	487	5 Q45012	Q45012 caenorhabdi
55	34	89.5	489	13 Q7ZXX3	Q7zxx3 xenopus lae
56	34	89.5	494	17 Q974C6	Q974c6 sulfolobus
57	34	89.5	508	3 Q9P7S7	Q9p7s7 schizosacch
58	34	89.5	508	17 Q97WR4	Q97wr4 sulfolobus
59	34	89.5	510	5 Q9U5W4	Q9u5w4 drosophila
60	34	89.5	511	5 Q9VM69	Q9vm69 drosophila
61	34	89.5	519	13 Q803P5	Q803p5 brachydanio
62	34	89.5	533	10 Q9PFT6	Q9pft6 arabidopsis
63	34	89.5	533	10 Q9MAB3	Q9mab3 arabidopsis
64	34	89.5	536	11 Q8C8Y7	Q8c8y7 mus musculu
65	34	89.5	550	10 Q65335	Q65335 pisum sativ
66	34	89.5	555	10 Q9AV96	Q9av96 nicotiana t
67	34	89.5	560	10 Q65334	Q65334 pisum sativ
68	34	89.5	597	3 Q8X066	Q8x066 neurospora
69	34	89.5	774	11 Q8C771	Q8c771 mus musculu
70	34	89.5	785	16 Q9CN05	Q9cn05 pasteurella
71	34	89.5	2659	5 Q9VB07	Q9vbu7 drosophila
72	33	86.8	129	5 Q8MS67	Q8ms67 drosophila
73	33	86.8	260	16 Q83250	Q83250 treponema p
74	33	86.8	283	16 Q91065	Q91065 pseudomonas
75	33	86.8	488	16 Q9A5V0	Q9a5v0 caulobacter
76	33	86.8	492	16 Q8YUZ1	Q8yuz1 anabaena sp
77	33	86.8	549	5 Q60963	Q60963 leishmania
78	33	86.8	582	16 Q89P35	Q89p35 bradyrhizob
79	33	86.8	779	16 Q8YX39	Q8yx39 anabaena sp
80	32	84.2	227	11 F70471	F70471 rattus norv
81	32	84.2	232	11 F70472	F70472 rattus norv
82	32	84.2	244	10 Q04953	Q04953 arabidopsis
83	32	84.2	274	16 Q7U099	Q7u099 mycobacteri
84	32	84.2	358	16 Q8XQ01	Q8xq01 ralsstonia s
85	32	84.2	377	16 Q985U9	Q985u9 rhizobium l
86	32	84.2	396	16 Q8ZNU7	Q8znj7 salmonella
87	32	84.2	396	16 Q8X5A2	Q8x5a2 escherichia
88	32	84.2	396	16 Q8FFS3	Q8ffs3 escherichia
89	32	84.2	396	16 Q8Z581	Q8z581 salmonella

90	32	84.2	396	16	Q83KE0	Q83ke0 shigella fl	163	30	78.9	495	16	Q8P6M5	Q8p6m5 xanthomonas
91	32	84.2	441	16	Q911Q7	Q911q7 pseudomonas	164	30	78.9	495	16	Q7WJW3	Q7Wjw3 bordetella
92	32	84.2	477	11	Q623P9	Q623p9 rattus norv	165	30	78.9	501	16	Q82D02	Q82d02 streptomyce
93	32	84.2	483	11	Q623P9	Q623p9 rattus norv	166	30	78.9	504	16	Q8PHX9	Q8phx9 xanthomonas
94	32	84.2	527	2	Q8GDM1	Q8gdm1 photorhabdu	167	30	78.9	504	16	Q8KZR2	Q8kzr2 streptomyce
95	32	84.2	699	16	Q82QU9	Q82qu9 streptomyce	168	30	78.9	539	10	Q8RUK2	Q8ruk2 oryza sativ
96	32	84.2	1181	2	Q8GDN1	Q8gdn1 photorhabdu	169	30	78.9	539	10	Q7XFF9	Q7Xff9 oryza sativ
97	31	81.6	53	16	Q99V62	Q99v62 staphylococ	170	30	78.9	541	10	Q8RUV6	Q8ruv6 oryza sativ
98	31	81.6	116	11	Q9D1A9	Q9d1a9 mus musculu	171	30	78.9	541	10	Q7XFF8	Q7Xff8 oryza sativ
99	31	81.6	148	16	Q9RDS5	Q9rds5 streptomyce	172	30	78.9	551	16	Q8AAL3	Q8aal3 bacteroides
100	31	81.6	148	16	Q827F7	Q827f7 streptomyce	173	30	78.9	551	4	Q13041	Q13041 homo sapien
101	31	81.6	168	16	Q7WDX6	Q7wxd6 bordetella	174	30	78.9	641	16	Q7UWE3	Q7uwe3 rhodopirell
102	31	81.6	168	16	Q7W8Z3	Q7w8z3 bordetella	175	30	78.9	655	16	Q89PK3	Q89pk3 bradyrhizob
103	31	81.6	168	16	Q7VXK4	Q7vxk4 bordetella	176	30	78.9	678	16	Q9A2E6	Q9a2e6 caulobacter
104	31	81.6	177	16	Q89XK3	Q89xk3 bradyrhizob	177	30	78.9	689	16	Q8NTZ7	Q8ntz7 corynebacte
105	31	81.6	209	11	Q9D9N2	Q9d9n2 mus musculu	178	30	78.9	715	16	Q8FUF9	Q8fuf9 corynebacte
106	31	81.6	209	11	Q80ZS5	Q80zs5 mus musculu	179	30	78.9	747	16	Q8P0D2	Q8p0d2 streptococ
107	31	81.6	217	16	Q8FC64	Q8fc64 escherichia	180	30	78.9	747	16	Q8K6Y6	Q8k6y6 streptococ
108	31	81.6	238	11	Q921D3	Q921d3 rattus norv	181	30	78.9	747	16	Q878X7	Q878x7 streptococ
109	31	81.6	241	16	Q82BG9	Q82bg9 streptomyce	182	30	78.9	791	5	Q9U185	Q9u185 leishmania
110	31	81.6	295	3	Q8J244	Q8j244 phaeosphaer	183	30	78.9	865	16	Q89BR1	Q89br1 bradyrhizob
111	31	81.6	300	3	Q8J240	Q8j240 mycosphaere	184	30	78.9	883	9	Q858N4	Q858n4 yersinia pe
112	31	81.6	384	16	Q7WJE4	Q7wje4 bordetella	185	30	78.9	1056	16	Q83116	Q83116 tropheryma
113	31	81.6	384	16	Q7WAA1	Q7waa1 bordetella	186	30	78.9	1066	16	Q83G52	Q83g52 tropheryma
114	31	81.6	403	16	Q7VYH3	Q7vyh3 bordetella	187	29	76.3	55	16	Q9AK97	Q9ak97 streptomyce
115	31	81.6	408	16	Q8VYD7	Q8vyd7 arabidopsis	188	29	76.3	60	5	Q9TXF9	Q9txf9 boophilus m
116	31	81.6	419	16	Q99WX6	Q99wx6 staphylococ	189	29	76.3	71	12	Q91C04	Q91c04 kaposi's sa
117	31	81.6	419	16	Q8EW7	Q8emt7 oceanobacil	190	29	76.3	71	12	Q91C20	Q91c20 kaposi's sa
118	31	81.6	419	16	Q829U2	Q829u2 bacillus ha	191	29	76.3	71	12	Q91BZ4	Q91bz4 kaposi's sa
119	31	81.6	427	17	Q82TY6	Q82ty6 pyrobaculum	192	29	76.3	71	12	Q91C23	Q91c23 kaposi's sa
120	31	81.6	433	10	Q8M1W1	Q8m1w1 arabidopsis	193	29	76.3	71	12	Q91112	Q91112 kaposi's sa
121	31	81.6	465	10	Q8S6U1	Q8s6u1 oryza sativ	194	29	76.3	71	12	Q91BZ6	Q91bz6 kaposi's sa
122	31	81.6	467	16	Q7UB31	Q7ub31 shigella fl	195	29	76.3	71	12	Q91BZ7	Q91bz7 kaposi's sa
123	31	81.6	475	16	Q8ZKV6	Q8zkv6 salmonella	196	29	76.3	71	12	Q91BZ8	Q91bz8 kaposi's sa
124	31	81.6	475	16	Q8XAW5	Q8xaw5 escherichia	197	29	76.3	71	12	Q91C03	Q91c03 kaposi's sa
125	31	81.6	475	16	Q8FBS1	Q8fbs1 escherichia	198	29	76.3	71	12	Q80PD7	Q80pd7 kaposi's sa
126	31	81.6	475	16	Q83P19	Q83p19 shigella fl	199	29	76.3	71	12	Q80PD6	Q80pd6 kaposi's sa
127	31	81.6	480	16	P74370	P74370 synchocyst	200	29	76.3	71	12	Q80PD5	Q80pd5 kaposi's sa
128	31	81.6	1004	16	Q8EKJ1	Q8ekj1 shewanella	201	29	76.3	101	2	Q8GAG3	Q8geg3 arthrobacte
129	31	81.6	1099	16	Q8E0Z1	Q8e0z1 salmonella	202	29	76.3	106	16	Q9PE50	Q9pe50 xyliella fas
130	31	81.6	1154	11	Q8CDS5	Q8cds5 mus musculu	203	29	76.3	108	17	Q9YAA3	Q9yaa3 aeropyrum p
131	31	81.6	1177	16	Q8ZB26	Q8zb26 yersinia pe	204	29	76.3	109	11	Q8C3A4	Q8c3a4 mus musculu
132	31	81.6	126	4	Q8N366	Q8n366 homo sapien	205	29	76.3	110	16	Q9JZV7	Q9jzv7 neisseria m
133	30	78.9	65	16	Q7URSO	Q7urso rhodopirell	206	29	76.3	110	16	Q8TUX0	Q8tux0 neisseria m
134	30	78.9	126	4	Q8NV44	Q8nv44 homo sapien	207	29	76.3	110	16	Q88H41	Q88h41 pseudomonas
135	30	78.9	144	16	Q07171	Q07171 mycobacteri	208	29	76.3	111	15	Q8SNM2	Q8snm2 human immun
136	30	78.9	144	16	Q7U2T3	Q7u2t3 mycobacteri	209	29	76.3	113	16	Q8ZF55	Q8zif55 yersinia pe
137	30	78.9	149	16	Q9CME2	Q9cme2 pasteurellia	210	29	76.3	113	16	Q8KFP1	Q8kfp1 chlorobium
138	30	78.9	156	16	Q7WBG6	Q7wbg6 bordetella	211	29	76.3	114	12	Q995C0	Q995c0 kaposi's sa
139	30	78.9	157	16	Q8CJY8	Q8cjy8 streptomyce	212	29	76.3	114	16	Q8D0B9	Q8d0b9 yersinia pe
140	30	78.9	158	16	Q8DGA3	Q8dga3 synchococ	213	29	76.3	115	12	Q995A1	Q995a1 kaposi's sa
141	30	78.9	161	16	Q9PCD6	Q9pcd6 xyliella fas	214	29	76.3	116	16	Q8VWD6	Q8vwd6 streptomyce
142	30	78.9	163	16	Q7WMY7	Q7wmv7 bordetella	215	29	76.3	118	12	Q995C2	Q995c2 kaposi's sa
143	30	78.9	183	10	Q7XTD4	Q7xtcd4 oryza sativ	216	29	76.3	121	12	Q995E3	Q995e3 kaposi's sa
144	30	78.9	189	16	Q9PDG8	Q9pdg8 xyliella fas	217	29	76.3	124	16	Q82G09	Q82g09 streptomyce
145	30	78.9	189	16	Q87DP2	Q87dp2 xyliella fas	218	29	76.3	125	16	Q87WC8	Q87wc8 pseudomonas
146	30	78.9	202	16	Q828I9	Q828i9 streptomyce	219	29	76.3	126	16	Q8YXT6	Q8yxt6 anabaena sp
147	30	78.9	227	16	Q81BF8	Q81bf8 bacillus ce	220	29	76.3	135	11	Q9CPW3	Q9cpw3 mus musculu
148	30	78.9	235	17	Q9V0T2	Q9v0t2 pyrococcus	221	29	76.3	136	16	Q824Z4	Q824z4 streptomyce
149	30	78.9	240	2	Q9S1T4	Q9s1t4 streptomyce	222	29	76.3	137	16	Q99Q89	Q99qe9 streptomyce
150	30	78.9	288	16	Q9XDW4	Q9xdw4 rhodopseudo	223	29	76.3	142	16	Q9RRN1	Q9rrn1 deinococcus
151	30	78.9	318	10	Q8GZB5	Q8gzb5 hordium vul	224	29	76.3	142	16	Q81Q42	Q81q42 bacillus an
152	30	78.9	336	16	Q7UZL1	Q7uzl1 prochloroco	225	29	76.3	147	11	Q81427	Q81427 mus musculu
153	30	78.9	343	16	Q7V987	Q7v987 prochloroco	226	29	76.3	148	16	Q8CHV4	Q8chv4 lactococcus
154	30	78.9	354	2	Q939S7	Q939s7 desulfitoba	227	29	76.3	150	16	Q81D46	Q81d46 bacillus ce
155	30	78.9	388	10	Q7TID8	Q7tid8 oryza sativ	228	29	76.3	153	16	Q8JMS1	Q8jms1 escherichia
156	30	78.9	398	17	Q9YE58	Q9ye58 aeropyrum p	229	29	76.3	154	16	Q8XGK7	Q8xgk7 escherichia
157	30	78.9	431	13	Q7ZVG7	Q7zvg7 brachydanio	230	29	76.3	154	16	Q8FJ13	Q8fj13 escherichia
158	30	78.9	441	16	Q7U4Q4	Q7u4q4 synchococ	231	29	76.3	158	16	Q98CT6	Q98ct6 rhizobium l
159	30	78.9	487	2	Q8VPW3	Q8vpw3 desulfitoba	232	29	76.3	161	15	Q73157	Q73157 human immun
160	30	78.9	487	2	Q8VPW3	Q8vpw3 desulfitoba	233	29	76.3	162	15	Q73155	Q73155 human immun
161	30	78.9	487	2	Q8VPW3	Q8vpw3 desulfitoba	234	29	76.3	165	16	Q8PPX7	Q8ppx7 xanthomonas
162	30	78.9	487	2	P72184	P72184 propionibac	235	29	76.3				

236	29	76.3	168	16	Q9CKT4	Q9cktt4 pasteurilla	309	29	76.3	284	12	Q91GV0	Q91gv0 kaposi's sa
237	29	76.3	170	2	Q8VMH4	Q8vmh4 pseudomonas	310	29	76.3	285	12	Q9WHC3	Q9whc3 kaposi's sa
238	29	76.3	171	16	Q7UYX1	Q7uyx1 rhodospirill	311	29	76.3	285	12	Q9WHH9	Q9whh9 kaposi's sa
239	29	76.3	172	5	Q7YWM0	Q7ywm0 caenorhabdi	312	29	76.3	286	16	Q24958	Q24958 helicobacte
240	29	76.3	172	16	Q8XB29	Q8xb29 escherichia	313	29	76.3	287	16	Q9PMH5	Q9pmh5 campylobact
241	29	76.3	172	16	Q9L088	Q9l088 streptomyc	314	29	76.3	288	4	Q96F74	Q96f74 homo sapien
242	29	76.3	172	16	Q82EUS	Q82eus streptomyc	315	29	76.3	288	16	Q8ZM00	Q8zm00 salmonella
243	29	76.3	172	16	Q83L23	Q83l23 shigella fl	316	29	76.3	288	16	Q8Z3P3	Q8z3p3 salmonella
244	29	76.3	173	4	Q9UKC5	Q9ukc5 homo sapien	317	29	76.3	288	16	Q81PI9	Q81pi9 bacillus an
245	29	76.3	174	16	Q8DSY9	Q8dsy9 streptococc	318	29	76.3	289	5	Q8SVS4	Q8svs4 encephalito
246	29	76.3	174	16	Q7WLC7	Q7wlc7 bordetella	319	29	76.3	289	12	Q9WH95	Q9wh95 kaposi's sa
247	29	76.3	174	16	Q7VX31	Q7vx31 bordetella	320	29	76.3	289	12	Q9DSB8	Q9dsb8 kaposi's sa
248	29	76.3	176	16	Q7WY77	Q7wy77 bordetella	321	29	76.3	289	12	Q9DSF5	Q9dsf5 kaposi's sa
249	29	76.3	180	12	Q7TF15	Q7tf15 rhesus cyto	322	29	76.3	289	12	Q9IC84	Q9ic84 kaposi's sa
250	29	76.3	180	16	Q82IM5	Q82im5 streptomyc	323	29	76.3	289	12	Q9DSOC	Q9dsoc kaposi's sa
251	29	76.3	185	16	Q9X853	Q9x853 streptomyc	324	29	76.3	289	12	Q9DH11	Q9dh11 kaposi's sa
252	29	76.3	193	16	Q9KJH5	Q9klh5 vibrio chol	325	29	76.3	289	12	Q9WHC1	Q9whc1 kaposi's sa
253	29	76.3	197	16	Q9RJF5	Q9rjp5 streptomyc	326	29	76.3	289	12	Q9DSE4	Q9dse4 kaposi's sa
254	29	76.3	198	10	Q7XZA0	Q7xza0 griffithsia	327	29	76.3	289	12	Q9DSD0	Q9dsd0 kaposi's sa
255	29	76.3	205	16	Q99V11	Q99v11 staphylococ	328	29	76.3	289	12	Q9WHB9	Q9whb9 kaposi's sa
256	29	76.3	205	16	Q8NX12	Q8nx12 staphylococ	329	29	76.3	289	12	Q9WHB9	Q9whb9 kaposi's sa
257	29	76.3	205	16	Q8CKX9	Q8ckx9 yersinia pe	330	29	76.3	289	12	P88970	P88970 kaposi's sa
258	29	76.3	208	11	Q8BTS0	Q8bts0 mus musculu	331	29	76.3	289	12	Q9IC87	Q9ic87 kaposi's sa
259	29	76.3	213	16	Q88P25	Q88pj5 pseudomonas	332	29	76.3	289	12	Q9WHH8	Q9whh8 kaposi's sa
260	29	76.3	218	5	Q88P25	Q88pj5 pseudomonas	333	29	76.3	289	12	Q9DSF1	Q9dsf1 kaposi's sa
261	29	76.3	227	12	Q9MHA1	Q9mha1 kaposi's sa	334	29	76.3	289	12	Q9WHC9	Q9whc9 kaposi's sa
262	29	76.3	228	9	Q7Y513	Q7y513 bacterioph	335	29	76.3	289	12	Q9WHC7	Q9whc7 kaposi's sa
263	29	76.3	228	12	Q9W9W9	Q9w9w9 kaposi's sa	336	29	76.3	289	12	Q9IC89	Q9ic89 kaposi's sa
264	29	76.3	229	12	Q9WHB1	Q9whb1 kaposi's sa	337	29	76.3	289	12	Q9IC89	Q9ic89 kaposi's sa
265	29	76.3	232	2	Q83Y77	Q83y77 streptococc	338	29	76.3	289	12	Q80IF6	Q80if6 kaposi's sa
266	29	76.3	234	13	Q9W6U0	Q9w6ue fugu rubrip	339	29	76.3	289	12	Q80IF4	Q80if4 kaposi's sa
267	29	76.3	235	12	Q9IC40	Q9ic40 kaposi's sa	340	29	76.3	289	12	Q80IF3	Q80if3 kaposi's sa
268	29	76.3	244	5	Q9VSG1	Q9vsg1 drosophila	341	29	76.3	289	12	Q80IF2	Q80if2 kaposi's sa
269	29	76.3	245	16	Q8PKX8	Q8pkx8 xanthomonas	342	29	76.3	289	12	Q80IF1	Q80if1 kaposi's sa
270	29	76.3	246	16	Q8P9E9	Q8p9e9 xanthomonas	343	29	76.3	289	12	Q80IE7	Q80ie7 kaposi's sa
271	29	76.3	249	16	Q892P2	Q892f2 clostridium	344	29	76.3	289	12	Q80IE5	Q80ie5 kaposi's sa
272	29	76.3	251	5	Q8IQZ3	Q8iqz3 drosophila	345	29	76.3	289	12	Q80IE2	Q80ie2 kaposi's sa
273	29	76.3	252	10	Q8LEV7	Q8lev7 arabidopsis	346	29	76.3	289	12	Q80IE1	Q80ie1 kaposi's sa
274	29	76.3	254	16	Q9XAK4	Q9xak4 streptomyc	347	29	76.3	289	12	Q80ID9	Q80id9 kaposi's sa
275	29	76.3	256	16	Q7WL93	Q7wl93 bordetella	348	29	76.3	289	12	Q80ID7	Q80id7 kaposi's sa
276	29	76.3	256	16	Q7WV33	Q7wv33 bordetella	349	29	76.3	289	12	Q80ID6	Q80id6 kaposi's sa
277	29	76.3	256	16	Q7VX00	Q7vx00 bordetella	350	29	76.3	289	12	Q80ID5	Q80id5 kaposi's sa
278	29	76.3	258	10	Q94IQ4	Q94iq4 brassica ca	351	29	76.3	289	12	Q80ID5	Q80id5 kaposi's sa
279	29	76.3	259	10	Q8LG70	Q8lg70 arabidopsis	352	29	76.3	289	12	Q805H3	Q805h3 kaposi's sa
280	29	76.3	259	10	Q9SG13	Q9sg13 arabidopsis	353	29	76.3	290	12	Q9DSF2	Q9dsf2 kaposi's sa
281	29	76.3	259	16	Q8ZPZ5	Q8zpz5 yersinia pe	354	29	76.3	290	12	Q9IC90	Q9ic90 kaposi's sa
282	29	76.3	260	16	Q8ZPZ5	Q8zpz5 yersinia pe	355	29	76.3	290	12	Q80IE6	Q80ie6 kaposi's sa
283	29	76.3	260	16	Q8G7Y0	Q8g7y0 bifidobacte	356	29	76.3	290	12	Q80IE4	Q80ie4 kaposi's sa
284	29	76.3	261	10	Q9FIN1	Q9fin1 arabidopsis	357	29	76.3	292	2	O87196	O87196 helicobacte
285	29	76.3	261	16	Q8ZPE9	Q8zpe9 salmonella	358	29	76.3	292	16	Q9ZWT3	Q9zwt3 helicobacte
286	29	76.3	261	16	Q8ZPZ5	Q8zpz5 salmonella	359	29	76.3	296	2	Q9RFV1	Q9rfv1 mycobacteri
287	29	76.3	263	16	Q9A2U8	Q9a2u8 caulobacter	360	29	76.3	300	3	Q8J250	Q8j250 phaeosphaer
288	29	76.3	263	16	Q9IC88	Q9ic88 kaposi's sa	361	29	76.3	300	16	Q9RS72	Q9rs72 deinococcus
289	29	76.3	266	12	Q9IC88	Q9ic88 kaposi's sa	362	29	76.3	300	16	Q7U8M3	Q7u8m3 synchococc
290	29	76.3	267	12	Q9WHA9	Q9wha9 kaposi's sa	363	29	76.3	302	3	Q8J252	Q8j252 phaeosphaer
291	29	76.3	270	10	Q84XU3	Q84xu3 elaeis guin	364	29	76.3	302	3	Q8J249	Q8j249 phaeosphaer
292	29	76.3	272	12	Q9IC30	Q9ic30 kaposi's sa	365	29	76.3	303	15	Q76196	Q76196 human immu
293	29	76.3	272	12	Q9IC28	Q9ic28 kaposi's sa	366	29	76.3	304	17	Q9YG97	Q9yg97 aeropyrum p
294	29	76.3	272	12	Q9WHA7	Q9wha7 kaposi's sa	367	29	76.3	306	16	Q81NB7	Q81nb7 bacillus an
295	29	76.3	273	16	Q95B75	Q95b7 kaposi's sa	368	29	76.3	310	16	Q8FS58	Q8fs58 corynebacte
296	29	76.3	273	16	Q88FW5	Q88fw5 pseudomonas	369	29	76.3	310	16	Q8FS58	Q8fs58 corynebacte
297	29	76.3	275	12	Q995A4	Q995a4 kaposi's sa	370	29	76.3	312	4	Q8NG32	Q8ng32 homo sapien
298	29	76.3	276	12	Q9IC26	Q9ic26 kaposi's sa	371	29	76.3	313	16	Q8ZRV4	Q8zrv4 streptomyc
299	29	76.3	276	12	Q9IC27	Q9ic27 kaposi's sa	372	29	76.3	314	5	Q9NE71	Q9ne71 leishmania
300	29	76.3	276	12	Q9IC31	Q9ic31 kaposi's sa	373	29	76.3	314	11	Q8VC25	Q8vc25 mus musculu
301	29	76.3	276	12	Q9WNS7	Q9wns7 kaposi's sa	374	29	76.3	316	5	O62517	O62517 caenorhabdi
302	29	76.3	277	2	Q9IC25	Q9ic25 kaposi's sa	375	29	76.3	318	3	Q8Z6H5	Q8z6h5 saccharomyc
303	29	76.3	277	2	Q5F722	Q5f722 rhizobium 1	376	29	76.3	322	5	Q8S2E2	Q8s2e2 drosophila
304	29	76.3	279	2	Q95036	Q95036 synchocyst	377	29	76.3	323	3	Q03373	Q03373 saccharomyc
305	29	76.3	279	12	Q955B6	Q955b6 kaposi's sa	378	29	76.3	324	2	Q8GMP5	Q8gmp5 aeromonas s
306	29	76.3	279	12	Q955A3	Q955a3 kaposi's sa	379	29	76.3	324	2	Q84BF0	Q84bf0 xanthomonas
307	29	76.3	280	12	Q9WHB2	Q9whb2 kaposi's sa	380	29	76.3	328	2	Q7X2E2	Q7x2e2 aeromonas s
308	29	76.3	281	16	Q81CF6	Q81cf6 bacillus ce	381	29	76.3	328	16	Q88PI4	Q88pi4 pseudomonas
										329	10	Q9ZTB8	Q9ztb8 zea mays (m

382	29	76.3	335	2	Q8GMN8	Q8gmns8 aeromonas s	455	29	76.3	516	16	Q7ULA0	Q7ula0 rhodopirell
383	29	76.3	336	10	Q7XWP9	Q7xwp9 oryza sativ	456	29	76.3	521	5	Q95XV8	Q95xv8 caenorhabdi
384	29	76.3	337	16	Q7V9L9	Q7v9l9 prochloroco	457	29	76.3	522	5	Q9BHW0	Q9bhw0 leishmania
385	29	76.3	338	16	Q8DII5	Q8di15 synecococc	458	29	76.3	527	5	Q9VGT1	Q9vgt1 drosophila
386	29	76.3	339	2	Q8LOV9	Q8lv9 oryza sativ	459	29	76.3	530	16	Q7U445	Q7u445 synecococc
387	29	76.3	340	7	Q9XR09	Q9xr09 escherichia	460	29	76.3	533	10	Q8RY80	Q8ry80 arabidopsis
388	29	76.3	341	16	Q8RI52	Q8ri52 pseudomonas	461	29	76.3	541	5	O45287	O45287 caenorhabdi
389	29	76.3	342	16	Q7UXR6	Q7uxr6 rhodopirell	462	29	76.3	548	16	Q8RAK2	Q8rak2 streptomyce
390	29	76.3	343	5	Q9CSB2	Q9csb2 mus musculu	463	29	76.3	551	4	Q9BR26	Q9br26 homo sapien
391	29	76.3	344	11	Q9VX04	Q9vx04 drosophila	464	29	76.3	563	5	Q9N3N7	Q9n3n7 caenorhabdi
392	29	76.3	345	16	Q8NSY1	Q8nsy1 corynebacte	465	29	76.3	563	5	Q9N3N7	Q9n3n7 caenorhabdi
393	29	76.3	346	16	Q8E796	Q8e796 streptomyce	466	29	76.3	567	16	Q8ZC73	Q8zct73 yersinia pe
394	29	76.3	347	17	Q8ZWC3	Q8zwc3 pyrobaculum	467	29	76.3	569	16	Q8FGJ7	Q8fgj7 escherichia
395	29	76.3	348	5	Q8GM12	Q8gm12 myxine glut	468	29	76.3	574	3	Q874L9	Q874l9 kluyveromyc
396	29	76.3	349	16	Q8A4W8	Q8a4w8 bacteroides	469	29	76.3	579	9	Q8W723	Q8w723 cyanophaga
397	29	76.3	350	16	Q8S9S6	Q8s9s6 bradyrhizob	470	29	76.3	579	3	Q875R3	Q875r3 saccharomyc
398	29	76.3	351	16	Q7WHL3	Q7whl3 bordetella	471	29	76.3	582	3	Q7Z877	Q7z877 candida tro
399	29	76.3	352	16	Q7W6N1	Q7w6n1 bordetella	472	29	76.3	583	16	Q9L1F7	Q9l1f7 streptomyce
400	29	76.3	353	16	Q7VXB4	Q7vxb4 bordetella	473	29	76.3	587	3	Q875N1	Q875n1 kluyveromyc
401	29	76.3	354	4	Q9NXS2	Q9nxs2 homo sapien	474	29	76.3	591	2	Q938U7	Q938u7 streptomyce
402	29	76.3	355	16	Q8D8Y3	Q8d8y3 mus musculu	475	29	76.3	591	2	Q938U7	Q938u7 streptomyce
403	29	76.3	356	16	Q82AX9	Q82ax9 streptomyce	476	29	76.3	593	16	Q9KZJ7	Q9kzj7 streptomyce
404	29	76.3	357	11	Q9D8P5	Q9d8p5 mus musculu	477	29	76.3	593	16	Q82MU0	Q82mu0 streptomyce
405	29	76.3	358	11	Q8BH73	Q8bh73 mus musculu	478	29	76.3	594	16	Q82KR5	Q82kr5 streptomyce
406	29	76.3	359	16	Q9RKQ3	Q9rkq3 streptomyce	479	29	76.3	598	16	Q9KS01	Q9ks01 vibrio chol
407	29	76.3	360	16	Q81ZM4	Q81zm4 streptomyce	480	29	76.3	610	17	Q971Q0	Q971q0 sulfolobus
408	29	76.3	361	16	P96842	P96842 mycobacteri	481	29	76.3	621	5	Q95PS1	Q95ps1 leishmania
409	29	76.3	362	16	Q7TW74	Q7tw74 mycobacteri	482	29	76.3	621	5	Q95PS1	Q95ps1 leishmania
410	29	76.3	363	16	Q8X0B6	Q8x0b6 neurospora	483	29	76.3	622	16	Q82J76	Q82j76 arabidopsis
411	29	76.3	364	17	Q8ZSN9	Q8zsn9 pyrobaculum	484	29	76.3	626	2	Q8RMT0	Q8rmt0 acinetobact
412	29	76.3	365	2	Q9AJU1	Q9aju1 streptomyce	485	29	76.3	629	10	O49702	O49702 arabidopsis
413	29	76.3	366	10	Q9C693	Q9c693 arabidopsis	486	29	76.3	633	16	P74015	P74015 synecocyst
414	29	76.3	367	2	P95556	P95556 pseudomonas	487	29	76.3	633	16	Q8PKH7	Q8pkh7 xanthomonas
415	29	76.3	368	16	Q7UIG3	Q7uig3 rhodopirell	488	29	76.3	647	5	Q8MSU3	Q8msu3 drosophila
416	29	76.3	369	16	Q9IOH6	Q9ioh6 pseudomonas	489	29	76.3	652	2	Q93TK2	Q93tk2 streptococc
417	29	76.3	370	3	O13982	O13982 schizosacch	490	29	76.3	660	5	Q17248	Q17248 boophilus m
418	29	76.3	371	16	Q8ZGM1	Q8zgm1 yersinia pe	491	29	76.3	666	16	Q880V5	Q880v5 pseudomonas
419	29	76.3	372	5	O45778	O45778 hepatitis b	492	29	76.3	668	11	Q92179	Q92179 mus musculu
420	29	76.3	373	16	Q8NNQ9	Q8nnq9 corynebacte	493	29	76.3	669	5	Q21476	Q21476 caenorhabdi
421	29	76.3	374	16	Q8UQX0	Q8uqx0 agrobacteri	494	29	76.3	684	16	Q87TE2	Q87te2 vibrio para
422	29	76.3	375	16	Q8UQX0	Q8uqx0 rhodopirell	495	29	76.3	687	9	Q9MC93	Q9mc93 bacterioph
423	29	76.3	376	16	Q93368	Q93368 caenorhabdi	496	29	76.3	687	9	Q9MC93	Q9mc93 bacterioph
424	29	76.3	377	5	Q93368	Q93368 caenorhabdi	497	29	76.3	687	9	Q9MC93	Q9mc93 bacterioph
425	29	76.3	378	16	Q93368	Q93368 caenorhabdi	498	29	76.3	687	9	Q9MC93	Q9mc93 bacterioph
426	29	76.3	379	16	Q93368	Q93368 caenorhabdi	499	29	76.3	687	9	Q9MC93	Q9mc93 bacterioph
427	29	76.3	380	16	Q93368	Q93368 caenorhabdi	500	29	76.3	687	9	Q9MC93	Q9mc93 bacterioph
428	29	76.3	381	16	Q93368	Q93368 caenorhabdi							
429	29	76.3	382	16	Q93368	Q93368 caenorhabdi							
430	29	76.3	383	16	Q93368	Q93368 caenorhabdi							
431	29	76.3	384	16	Q93368	Q93368 caenorhabdi							
432	29	76.3	385	16	Q93368	Q93368 caenorhabdi							
433	29	76.3	386	16	Q93368	Q93368 caenorhabdi							
434	29	76.3	387	16	Q93368	Q93368 caenorhabdi							
435	29	76.3	388	16	Q93368	Q93368 caenorhabdi							
436	29	76.3	389	16	Q93368	Q93368 caenorhabdi							
437	29	76.3	390	16	Q93368	Q93368 caenorhabdi							
438	29	76.3	391	16	Q93368	Q93368 caenorhabdi							
439	29	76.3	392	16	Q93368	Q93368 caenorhabdi							
440	29	76.3	393	16	Q93368	Q93368 caenorhabdi							
441	29	76.3	394	16	Q93368	Q93368 caenorhabdi							
442	29	76.3	395	16	Q93368	Q93368 caenorhabdi							
443	29	76.3	396	16	Q93368	Q93368 caenorhabdi							
444	29	76.3	397	16	Q93368	Q93368 caenorhabdi							
445	29	76.3	398	16	Q93368	Q93368 caenorhabdi							
446	29	76.3	399	16	Q93368	Q93368 caenorhabdi							
447	29	76.3	400	16	Q93368	Q93368 caenorhabdi							
448	29	76.3	401	16	Q93368	Q93368 caenorhabdi							
449	29	76.3	402	16	Q93368	Q93368 caenorhabdi							
450	29	76.3	403	16	Q93368	Q93368 caenorhabdi							
451	29	76.3	404	16	Q93368	Q93368 caenorhabdi							
452	29	76.3	405	16	Q93368	Q93368 caenorhabdi							
453	29	76.3	406	16	Q93368	Q93368 caenorhabdi							
454	29	76.3	407	16	Q93368	Q93368 caenorhabdi							

ALIGNMENTS

RESULT 1

Q8GU78	ID	Q8GU78	PRELIMINARY;	PRT;	302 AA.
AC	Q8GU78;				
DT	01-MAR-2003	(T-EMBLrel. 23, Created)			
DT	01-MAR-2003	(T-EMBLrel. 23, Last sequence update)			
DT	01-JUN-2003	(T-EMBLrel. 24, Last annotation update)			
DE	Putative methyltransferase.				
GN	MTF4.				
OS	Mycobacterium smegmatis.				
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
OC	Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.				
OX	NCBI_TaxID=1772;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=2255602; PubMed=12368441;				
RA	Jeewanarajah D., Patterson J.H., McConville M.J., Billman-Jacobe H.;				
RT	"Modification of glycopeptidolipids by an O-methyltransferase of				
RT	Mycobacterium smegmatis."				
RL	Microbiology 148:3079-3087(2002).				
DR	EMBL; AY138899; AAN28690.1; "				
DR	GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . ; IEA.				

```

DR GO: GO:0016740; F:transferase activity; IEA.
DR InterPro: IPR000051; SAM bind.
KW Transferase; Methyltransferase.
SQ SEQUENCE 302 AA; 34721 MW; 09D6FA289C33C90E CRC64;

Query Match          94.7%; Score 36; DB 2; Length 302;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 10 WASWHF 15

RESULT 2
ID O59123 PRELIMINARY; PRT; 241 AA.
AC O59123;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PH1454.
GN PH1454.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yanamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000006; BAA30561.1; -.
DR PIR; A71020; A71020.
DR InterPro: IPR007182; MnhB.
DR Pfam; PF04039; MnhB; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 241 AA; 26494 MW; A095C10EAE70818 CRC64;

Query Match          92.1%; Score 35; DB 17; Length 241;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 193 WQWHF 198

RESULT 3
ID Q9FIN2 PRELIMINARY; PRT; 256 AA.
AC Q9FIN2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 6-phosphogluconolactonase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=99156233; PubMed=10048489;

RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5, VIII.
RT Sequence features of the regions of 1,081,958 bp covered by seventeen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:379-391(1998).
DR EMBL; AB016884; BAB1233.1; -.
DR GO: GO:0017057; P:6-phosphogluconolactonase activity; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR GO: GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro: IPR006148; Gluc_gal_isom.
DR InterPro: IPR005900; Phosphogluconlac.
DR Pfam; PF01182; Glucosamine iso; 1.
DR TIGRfams; TIGR01198; pgl; 1.
SQ SEQUENCE 256 AA; 28034 MW; 21E7046ECFDC72AF CRC64;

Query Match          92.1%; Score 35; DB 10; Length 256;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 72 WARWHF 77

RESULT 4
ID Q8NLA5 PRELIMINARY; PRT; 258 AA.
AC Q8NLA5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Uncharacterized ACR.
GN CGL3043.
OS Corynebacterium glutamicum (Brevibacterium flavum)
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005283; BAC00437.1; -.
KW Complete proteome.
SQ SEQUENCE 258 AA; 28686 MW; 871724DB2B3573C4 CRC64;

Query Match          92.1%; Score 35; DB 16; Length 258;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WXXWHF 6
Db 16 WHAWHF 21

RESULT 5
ID Q8PUY8 PRELIMINARY; PRT; 276 AA.
AC Q8PUY8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein MW2190.
GN MW2190.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;

```

```

RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias E., Henne A., Wierzer A., Baumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosaerina mazei: evidence for lateral gene
RL transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL: AE013459; AM31886.1; -.
DR InterPro: IPR003675; Abi.
DR Pfam: PF02517; Abi; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 276 AA; 31508 MW; 7F1BDB00D8D27521 CRC64;

Query Match 92.1%; Score 35; DB 17; Length 276;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 168 WSLWHF 173

RESULT 6
ID Q8TRL3 PRELIMINARY; PRT; 289 AA.
AC Q8TRL3
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein M1162.
GN M1162.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atncor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Pritchett M., Sowers K.R., Jing H., Macario A.J.L., Paulsen I.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL: AE010783; AM04583.1; -.
DR InterPro: IPR003675; Abi.
DR Pfam: PF02517; Abi; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 289 AA; 33289 MW; E422170BEA52B866 CRC64;

Query Match 92.1%; Score 35; DB 17; Length 289;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 168 WSLWHF 173

RESULT 7
Q9L120 PRELIMINARY; PRT; 314 AA.
ID Q9L120
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

```

```

AC Q9L120;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative hydrolase.
GN SC07440 OR SC6D11.36C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AU939131; CAB76359.1; -.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR GO: GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro: IPR000073; A/b hydrolase.
DR InterPro: IPR003089; AB_hydrolase.
DR InterPro: IPR000639; Epox_hydrolase.
DR InterPro: IPR000379; Ser_estr.
DR Pfam: PF00561; abhydrolase; 1.
DR PRINTS: PR00111; ABHYDROLASE.
DR PRINTS: PR00412; EPOXYDRASE.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 314 AA; 33938 MW; B302100E5CE852C7 CRC64;

Query Match 92.1%; Score 35; DB 16; Length 314;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
Db 168 WXTWHF 173

RESULT 8
Q8LD77 PRELIMINARY; PRT; 325 AA.
ID Q8LD77
AC Q8LD77
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

```

01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 6-phosphogluconolactonase-like protein.
 Arabidopsis thaliana (Mouse-ear cress).
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; rosids;
 eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 [1]
 SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RA "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0 (2002).
 [2]
 SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RA "Full-length cDNA from Arabidopsis thaliana";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY086161; AAM63366.1; -;
 DR GO; GO:0017057; F:6-phosphogluconolactonase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006148; Gluc_gal_isom.
 DR InterPro; IPR005900; Phosphogluconlac.
 DR Pfam; PF01182; Glucosamine iso; 1.
 DR TIGRFAMs; TIGR01198; pgl; 1.
 SQ SEQUENCE 325 AA; 35592 MW; C51D8544114943D2 CRC64;

Query Match 92.1%; Score 35; DB 10; Length 325;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 DB 141 WARWHF 146

RESULT 9
 Q84WW2 PRELIMINARY; PRT; 325 AA.
 ID Q84WW2
 AC Q84WW2
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative 6-phosphogluconolactonase.
 GN AT5G24400
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 [1]
 SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
 RA Onodera C.S., Quach H.H., Tang C.C., Toriumi M., Wong C., Wu H.C.,
 RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
 RA Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RA "Arabidopsis Open Reading Frame (ORF) Clones";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT001923; AAN71922.1; -;
 DR GO; GO:0017057; F:6-phosphogluconolactonase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006148; Gluc_gal_isom.
 DR InterPro; IPR005900; Phosphogluconlac.
 DR Pfam; PF01182; Glucosamine iso; 1.
 DR TIGRFAMs; TIGR01198; pgl; 1.
 SQ SEQUENCE 325 AA; 35644 MW; 0F6C9E9F9C073DC CRC64;

Query Match 92.1%; Score 35; DB 10; Length 325;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 DB 141 WARWHF 146

Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WXXWHF 6
 DB 141 WARWHF 146

RESULT 10
 Q8G242 PRELIMINARY; PRT; 377 AA.
 ID Q8G242
 AC Q8G242
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Modification methylase Babi.
 GN BABI OR BR0491.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 animal and plant pathogens and symbionts.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; AE014358; AAN29434.1; -;
 DR TIGR; BR0491; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
 DR GO; GO:0006306; P:DNA methylation; IEA.
 DR InterPro; IPR002295; D21N6 mtfrase.
 DR InterPro; IPR001091; Met trans CNA.
 DR InterPro; IPR002941; N6/N4 Mase.
 DR InterPro; IPR002052; N6 Mase.
 DR InterPro; IPR000051; SAM bind.
 DR Pfam; PF01555; N6_N4 Mase; 1.
 DR PRINTS; PR00506; D21N6MTFRASE.
 DR PRINTS; PR00508; S21N4MTFRASE.
 DR PROSITE; PS00092; N6_MTASE; 1.
 DR Methyltransferase; Complete proteome.
 KW Methyltransferase; Complete proteome.
 SQ SEQUENCE 377 AA; 42188 MW; 7E9D36B9F52BA83D CRC64;

Query Match 92.1%; Score 35; DB 16; Length 377;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 DB 347 WTWWHF 352

Query Match 92.1%; Score 35; DB 16; Length 377;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WXXWHF 6
 DB 347 WTWWHF 352

RESULT 11
 Q93TQ4 PRELIMINARY; PRT; 381 AA.
 ID Q93TQ4
 AC Q93TQ4
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Call cycle-regulated methyltransferase CorM.
 OS Agrobacterium tumefaciens.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OX NCBI_TaxID=358;